의학박사 학위논문

한국인 급성 골수성 백혈병의 돌연변이 분석과 그 임상적 의미에 관한 연구

Somatic mutations and their clinical implications in Korean acute myeloid leukemia patients

2018년 2월

서울대학교 대학원

의학과 분자종양의학 전공

고 영 일

한국인 급성 골수성 백혈병의 돌연변이 분석과 그 임상적 의미에 관한 연구

Somatic mutations and their clinical implications in Korean acute myeloid leukemia patients

지도교수 허 대 석

이 논문을 의학박사 학위논문으로 제출함

2017 년 10 월

서울대학교 대학원

의학과 분자종양의학 전공

고 영 일

고영일의 의학박사 학위논문을 인준함

2017 년 12 월

위 원 장 이 동 순 (인)

부 위 원 장 윤 성 수 (인)

위 원 최 무 림 (인)

위 원 김 형 래 (인)

위 원 허 대 석 (인)

Abstract

Background

Although, the mutational landscape of acute myeloid leukemia (AML) has been revealed by means of next-generation sequencing (NGS), low frequency mutation is yet to be discovered. In the present study, we performed NGS with bone marrow specimens to identify novel oncogenic mutations in AML at various clinical stages.

Materials and Methods

We performed whole exome sequencing (WES) of 53 AML samples, followed by targeted resequencing of 389 AML samples. Targeted resequencing was performed for all non-synonymous variants revealed by WES. Samples used for targeted resequencing included samples at diagnosis (n=155), at relapse/persistence (n=63) and at complete remission (n=87). The results were filtered using Exome Aggregation Consortium, 1000 Genome Database, Korean Variant Archive (KOVA) and The Cancer Genome Atlas.

Results

We could identify 590 loci that is potentially related to AML with WES. Taregted resequencing of 590 loci revealed 29 loci is recurrently found in AML samples. Mutation frequency of well known oncogenic variants including *IDH2 p.R172K, IDH2 p.R140H, NRAS p.G13D, NRAS p.G12D, and DNMT3A p.R693C* did not differ from the frequency of TCGA database. Whole Seven variants (*PCDHA1 p.L666F*, *SERAC1 p.R645C*, *C7orf50 p.H113Y*, *TRRAP p.S722F*, *FAM178 p.T105M*, *GATA2 p.R330L*, and *PIWIL3 p.N872S*) were found to be putative novel oncogenic mutations. Results of an in silico prediction tool suggested that *GATA2 p.R330L* and *TRRAP p.S722F* were highly likely to be functional oncogenic mutations. The oncogenic mutation *TRRAP p.S722F*, which had been described previously in melanoma, was found only in acute promyelocytic leukemia (APL) samples. Sequencing of specimens from an additional 24 patients with APL revealed a frequency of 12.8% for *TRRAP p.S722F*.

Conclusion

We found 7 novel mutations in AML, including *GATA2 p.R330L* and *TRRAP p.S722F*, which are putative oncogenic driver. *TRRAP p.S722F*

especially appears to be a partner mutation of the PML-RARα fusion in more than 10% of patients with APL.

Keywords: Acute myeloid leukemia, TRRAP, next generation sequencing, acute promyelocytic leukemia

Student Number: 2010-30558

Contents

Introduction

Acute myeloid leukemia (AML) is characterized by infiltration of the bone marrow, blood, and other tissues by proliferative, clonal, abnormally differentiated cells of the hematopoietic system.[1] In young patients with AML, the rate of complete remission (CR) may exceed 80%, with a 5-year overall survival (OS) of 40%.[2] In patients who are older, treatment with hypomethylating agents has improved median and short-term OS, but this has not translated into improved cure rates, which remain disappointingly low.[2] Therefore, AML has been the subject of numerous studies that focus on unraveling the pathogenesis and molecular heterogeneity of this disease with the aims of better understanding and treating AML.

Tremendous progress in next generation sequencing (NGS) technologies has changed our understanding of the genomic landscape of AML. Authors who sequenced DNA samples of 1540 patients with AML characterized relatively common mutations associated with this disease in terms of detailed frequencies and clinical importance.[3] In AML, mutations are known to occur in *FLT3, NPM1, DNMT3A, CEBPA, c-KIT, IDH2, TP53, SRSF2, NRAS, RUNX1, MLL, STAG2,* and *ASXL1*. Despite an extensive characterization of genetic variation in AML, additional oncogenic mutations—particularly those that occur less commonly—likely remain to

be identified. In the most recent study of the AML genetic landscape, ≥2 known oncogenic mutations were found in 86% of patients,[3] suggesting that 14% of patients with AML harbor mutations that require further molecular characterization.

NGS technology also has enabled an unprecedented overview of molecular heterogeneity and clonal evolution in AML.[4] The diversity and relative dominance of subclones of AML vary throughout the course of disease.[5, 6] Longitudinal sequencing studies of diagnostic and relapse specimens have shown heterogeneous patterns of relapse, with relapse occurring from expansion of major or minor clones present at diagnosis or from novel clones that share an ancestral relationship to the diagnostic clone.[6, 7] Biological interpretation of these longitudinal genomic studies has included the assignment of functional significance to specific mutations. For instance, mutations that persist after treatment failure likely constitute critical molecular drivers of leukemia development and relapse after chemotherapy. Additional research that involves AML specimens from diverse clinical stages of disease is essential to elucidate the biological significance of specific mutations.

We comprehensively analyzed novel mutations identified from whole exome sequencing (WES) in patients with AML, and we conducted

longitudinal sequencing of various disease phases. The aim of the current study was to reveal genetic changes with relatively low frequencies by targeting genetic changes at "recurrent loci," as determined by WES. We sequenced specimens from various stages of AML, including diagnosis, relapse/persistence, and CR to determine the biological relevance of specific genetic changes. We also correlated longitudinal changes in molecular profiles with patient outcomes. Deciphering cancer in the context of genomics is valuable for development of therapeutics that target frequent oncogenic mutations. In fact, clinical outcomes of patients with mutations in *FLT3*[8] and *IDH*[9] have been improved with targeted agents. Our findings, described herein, reveal additional oncogenic mutations in AML.

Materials and methods

Patients and samples

A total of 305 bone marrow samples and 84 saliva samples were obtained from adult patients with AML. All patients gave informed consent, and patients had various disease statuses at time of specimen collection. Specifically, 155 bone marrow specimens were obtained at time of diagnosis, 13 in persistence, 87 in CR after induction chemotherapy, and 50 in relapse. Serum was removed by centrifugation at 2000 rpm for 5 minutes, and samples were stored at −20°C. The 84 saliva samples were collected at time of CR and were stored in saliva preservation tubes (Norgen Biotek, Canada). Saliva samples served as a germline control.

Diagnosis of AML was based on criteria of the World Health Organization. Bone marrow smears were examined by means of May-Giemsa staining, and cytogenetic abnormalities were detected by conventional G-banding.[10] Patients were stratified into 3 risk groups based on results of cytogenetic and molecular analyses of bone marrow samples, according to the refined criteria of the Medical Research Council (MRC).[11] Data regarding patient demographics, response to chemotherapy, relapse-free survival (RFS), and OS were obtained by review of medical records. CR was defined as

normocellular bone marrow containing <5% blasts and showing evidence of normal maturation of other marrow elements, according to criteria developed by the International Working Group.[10, 12, 13] OS was defined as the duration from diagnosis to death from any cause, and RFS was defined as the time from diagnosis until relapse or death from any cause. Living patients were censored at date of last contact.

DNA preparation

Bone marrow samples were stored at -20℃ after removing serum using centrifuge at 2000 rpm for 5 minutes. Saliva sample were collected at the time of complete remission, and stored in saliva preserve tube (Norgen Biotek, Canada) following the manufacturer's recommendation. Tumor and germline DNA were extracted from bone marrow and saliva samples for whole exome sequencing (WES) and targeted re-sequencing. QuickGene DNA whole blood kit S (KURABO INDUSTRY, Japan) was used to extract DNA according to the manufacturer's recommendations.

Whole exome sequencing

Solexa sequencing technology platform (HiSeq2000, Illumina, SanDiego, CA, USA) was used for preparing and sequencing exomes following the manufacturer's instructions. For each sample, 3μg of genomic DNA was randomly sheared using Covaris System (Thermo Fisher Scientific, USA to generate about 150-bp inserts. The fragmented DNA was end-repaired using T4 DNA polymerase and Klenow polymerase, and Illumina paired-end adaptor-oligonucleotides were ligated to the sticky ends. The ligation mixture was analyzed by electrophoresis on an agarose gel. The length of sliced and purified fragments was 200-250 bp. Purified DNA library was hybridized with SureSelect Human All Exons V3 and V5 probes set (Agilent, Santa Clara, CA, USA) to capture 50Mb targeted exons following manufacture's instruction. We prepared the HiSeq2000 paired-end flowcell using captured exome library, according to the manufacturer's protocol. Clusters of PCR colonies were then sequenced on the HiSeq2000 platform using recommended protocols from the manufacturer.

Alignment of FASTQ file

FASTQ files were aligned to human reference (human_g1k_v 37.fasta) using the Burrows-Wheeler aligner (BWA-0.7.5) to make SAM (sequence alignment map) file [14]. SortSam in Picard-tools-1.68 was used to convert

to BAM (binary alignment map) file and sort by chromosome. BAM file went through a PCR duplicate marking process, which enables the Genome Analysis Toolkit (GATK-1.6.5) to ignore duplicates in subsequent processing [15]. Local realignment was performed prior to recalibration, which gives the most accurate quality scores for each sample. Recalibration was performed to increase recalibration accuracy. All of processes were set as a default.

Calling method using Adiscan which is in-house program

Each locus or chromosomal positions consists of two alleles (allele A, allele B) from parents which are equally balanced to 50:50. However, in some cases, allele fractions are not balanced, and thereby assumed to be calling errors or false positives. In fact, 10 to 20 percent of allele fractions from NGS data have extreme difficulties for making any decision for the genotype assignment because of the unbalanced feature (Table 1).

Supposed that a total depth of both of allele in a certain locus on a chromosome is near 50, PAF (pair allele fraction) type 1 stands for one of two samples not sequenced, so that their depths of one of two samples are zero. The opposite directions stand for homo to hetero and homo to

alternative homo, facing each other (large to small allele depths vs. small to large allele depths) (PAF types 2 to 5). The same directions stand for homo to heterozygote (large to small vs. large to small), and they are in the same direction (PAF types 6 to 8).

ADISCAN (allelic depth and imbalance scanning) is developed in such a way to meet the hypothesis that the pair allele fractions in a locus on normal genome and another one in cancer genome are unbiased and balanced. Therefore, the distance between paired allele fractions from normal and tumor samples can be defined as a tangent function as the bellows:

$$\mathrm{x(or\ y)} = \frac{\mathrm{small(A,B)}}{\mathrm{large(A,B)}}$$

$$\mathrm{t} = \tan^{-1}(\frac{x}{y})\ \tan^{-1}\left[\frac{\max(y)}{\max(x)}\right]\ \ where, y > x\ and\ \max(x) > \max(y)$$

where A and B stand for depths of allele A and B, and x and y are binning numbers (ranged 1 to 20) of paired allele fractions from normal and tumor samples, respectively. The small(A,B) and large(A,B) are simple functions that return smaller and larger value of (A,B), respectively. The max(x or y) returns the largest value in x or y. Then, the Adiscan score can be defined as:

$$\text{Adiscan score} = \log(40\cdot\mathrm{t})\cdot w_1 - \log[small(A,B,A',B')]\cdot w_2\cdot w_3$$

where w_1, w_2 and w_3 are weights defined a dependent weight on PAF types and their distance, sequencing errors and a simple constant, respectively.

Targeted resequencing

For the purpose of identifying loci for targeted resequencing, a subset of bone marrow samples (n = 53) were processed by WES and analyzed with in-house method Adiscan. This search yielded 590 loci. A total of 389 study samples were sequenced for this panel. A QuickGene DNA whole blood kit S (Kurabo Industry, Japan) subsequently was used to extract DNA from bone marrow and saliva samples. Purified DNA was obtained with a Qiagen DNeasy kit (Qiagen, Valencia, CA) and was assayed and quality-controlled using a Qubit 2.0 Fluorimeter (Life Technologies, Carlsbad, CA). Targeted sequencing then was performed by means of an Ion AmpliSeq custom panel (Life Technologies). To adequately detect long tail somatic mutations, all nonsynonymous single nucleotide variants (SNVs) were resequenced (including singlets). Library construction was performed using an Ion AmpliSeq Library Kit 2.0 (Life Technologies). Library templates were prepared and barcoded for sequencing with an Ion OneTouch System. Barcoded samples were multiplexed per Ion PI Chip (Life Technologies)

and were sequenced with an Ion Proton Sequencer System (Life Technologies). Sequencing reads were processed using Ion Torrent Suite software, version 4.2.1 (Life Technologies). De-multiplexed samples were assessed for sequencing quality, and high-quality reads were mapped to the complete hg19 human genome (UCSC version, February 2009). Variant discovery was carried out with Torrent Variant Caller, version 4.2 (Life Technologies), a software plugin for the Ion Torrent Suite. To eliminate errors in base calling, filtering steps were applied to generate final variant calling. The first filter was set at an average depth of total coverage of >30, each variant coverage of >10, a variant frequency of each sample >0.1, and $P < 0.01$. The second filter excluded putative strand-specific errors, such as a mutation detected in either the plus or minus strand of DNA but not in both strands.

Statistical analysis

Clinical characteristics of samples from 5 different clinical status were compared by analysis of variance (ANOVA) for continuous variables or by the Kruskal-Wallis method for ordinal variables. Differences in continuous variables between 2 groups were analyzed by unpaired *t* test or Mann-Whitney U test, as appropriate. A survival analysis was performed using the

Kaplan-Meier method, and results were compared with a log-rank test. Univariate and multivariate analyses with the Cox proportional hazard regression model were applied to determine hazard ratios (HRs) for specific variables with respect to OS and PFS. For all statistical analyses, two-sided *P*-values < 0.05 were considered statistically significant. All statistical analyses were carried out using STATA, version 12 (StataCorp LP, College Station, TX) and R, version 3.2.3 (http://www.r-project.org).

Ethical considerations

All patients provide their written informed consent to use their samples for banking and molecular analysis. The protocol was approved by the Seoul National University Hospital Institutional Review Board (IRB approval number: 1201-099-396). This study was conducted in accordance with Declaration of Helsinki provisions.

Results

Clinical characteristics of samples

The clinical characteristics of 389 samples are summarized in Table 2. More than half of the patients (59.6%) were women, and the patients' mean age was 52 years. The majority of the patients had favorable (24.9%) or intermediate (62.5%) MRC risk. The 3 most frequent FAB (i.e., French-America-British) classifications were M2 (29.8%), M3 (12.6%), and M4 (15.7%). A total of 117 patients (30.1%) underwent allogeneic stem cell transplantation. The male-to-female ratio, the proportion of secondary AML, and the FAB classification were comparable in all sample types.

Identification of novel SNVs

Of the 590 variants involving 480 genes (Table 3), 191 variants with <1% frequency in all samples were excluded (Figure 1). Twenty-seven variants with ≥10% frequency in saliva were regarded as germline variants. Filtering the variants via the 1000 Genomes Project and the Exome Aggregation Consortium (ExAC) databases [16] yielded 1 putative novel germline mutation: *TMPRSS13 p.Q78R*. A total of 372 variants with ≥1% frequency in bone marrow samples at time of diagnosis were considered somatic

variants. Of these, 360 variants were excluded, as outlined in Figure 1. The remaining 12 variants, involving 10 genes, were regarded as putative novel somatic mutations (Table 4). Five of these (*IDH2 p.R140Q, NRAS p.G13D, DNMT3A p.R693C, NRAS p.G12D,* and *IDH2 p.R172K*) were specified in The Cancer Genome Atlas (TCGA) (Figure 2). To our knowledge, the other 7 mutations (*TRRAP p.S722F, GATA2 p.R330L, FAM178B P.T105M, PCDHA1 p.L666F, C7orf50 p.H113Y, SERAC1 p.R645C,* and *PIWIL3 p.N872S*) have not been described previously. Korean Variant Archive (KOVA) [17] was used to assess possible Korean specific polymorphism in these genes.

Longitudinal distribution of the 12 identified variants

At time of diagnosis, although the 12 somatic variants were not identified in 102 of 155 patients (65.8%), the other 53 patients (34.2%) harbored at least 1 of these variants. Specifically, 44 patients carried 1 mutation, 8 patients had 2 mutations, and 1 patient carried 3 mutations. The mean number of mutations per patient (± standard deviation [SD]) was 0.40 ± 0.62 overall. This value was significantly higher in patients aged >52 years (0.48 ± 0.64) than in those aged ≤52 years (0.29 ± 0.57) (P = 0.028). Of samples obtained at time of diagnosis, 0.3% harbored ≥1 of the 12 somatic variants

(Figure 3). Specimens from patients in persistence appeared to be more likely to harbor these mutations, but the different was not significant (P = 0.613). Samples from patients in CR had a mutational proportion of 0.2%, which was significantly lower than samples from diagnosis (P = 0.035). Specimens from patients in relapse had a mutational proportion of 0.7%, which was significantly higher than CR samples (P = 0.009).

Among the 12 identified variants, *TRRAP p.S722F,* and *GATA2 p.R330L* were found only in diagnosis samples (Figure 4A). The remaining 10 mutations were found in both relapse and diagnosis samples, of which 6 mutations (*NRAS p.G12D, PCDHA1 p.L666F, C7orf50 p.H113Y, DNMT3A p.R693C, SERAC1 p.R645C*, and *PIWIL3 p.N872S*) were identified in CR samples as well. *IDH2 p.R140Q* was not found in CR samples but was found in persistence samples. The other 3 mutations (*NRAS p.G13D, IDH2 p.R172K*, and *FAM178B p.T105M*) occurred only in diagnosis and relapse samples.

The circos plot in Figure 4B demonstrates several patterns of correlation among the 12 variants. A significant positive correlation was observed in the following mutation pairs: *IDH2 p.R172K* and *FLT3 p.D835V*; *NRAS p.G12D* and *TRRAP p.S722F*, *GATA2 p.R330L*; *PCDHA1 p.L666F* and *NRAS p.G13D*; *PIWIL3 p.N872S* and *FLT3 p.D835V*, *GATA2 p.R330L* and

IDH2 p.R172K; *IDH2 p.R140Q* and *C7orf50 p.H113Y*, and *DNMT3A p.R693C* and *SERAC1 p.R645C* (Figure 4C).

Putative oncogenic mutations

The functional consequences of each variant were assessed by means of the Sorting Intolerant From Tolerant (SIFT) and Combined Annotation Dependent Depletion (CADD) algorithm (Table 4).[18, 19] Excluding variants commonly identified in the TCGA database, variants with a SIFT score ≤0.05 were *GATA2 p.R330L*, *PCDHA1 p.L666F, C7orf50 p.H113Y*, and *TRRAP p.S722F*. Among these, *TRRAP p.S722F* and *GATA2 p.R330L* had the highest CADD scores (4.926 and 5.528, respectively) and were not found in germline samples (Figure 4A), suggesting that these are purely somatic oncogenic mutations.

Association of somatic mutations with survival outcomes

At diagnosis, univariate analysis revealed 3 risk factors for RFS: secondary AML (HR 2.28; 95% confidence interval [CI] 0.99-5.26; *P*=0.054), NRAS p.G12D (HR 21.49; 95% CI 1.95-238.07; *P*=0.012), and PCDHA1 p.L666F (HR 6.38; 95% CI 1.69-24.13; *P*=0.006) (Table 5). For OS in patients with

samples obtained at diagnosis, age (>52 years old; HR, 1.92; 95% CI, 1.09-3.41; P = 0.025), male sex (HR, 1.98; 95% CI, 1.08-3.66; P = 0.028), unfavorable MRC risk group (HR, 4.07; 95% CI, 1.92-8.62; P < 0.001), and secondary AML (HR, 2.33; 95% CI, 1.26-4.32; P = 0.007) were identified as unfavorable prognostic factors. For patients in CR, no prognostic factor was associated with RFS. In contrast, age (>52 years old; HR, 8.35; 95% CI, 1.89-36.90; P = 0.005), poor MRC risk group (HR, 5.50; 95% CI, 1.74-17.39; P = 0.003), and presence of *NRAS p.G12D* (HR, 20.55; 95% CI, 2.30-183.83; P = 0.007) were found to be unfavorable risk factors for OS.

The independent prognostic significance of several somatic mutations was assessed using multivariate Cox proportional hazard regression with adjusting factors that were found to be associated with survival outcomes in univariate analysis. The HR and its 95% CI were summarized using forest plots (Figure 5). *NRAS p.G12D* (HR, 14.96; 95% CI, 1.17-191.83; P = 0.038) and presence of *PCDHA1 p.L666F* at diagnosis (HR, 8.13; 95% CI, 2.06-32.14; P = 0.003) were associated significantly with unfavorable RFS. Patients with *DNMT3A p.R693C* at diagnosis tended to have poorer RFS, although the difference was not statistically significant (HR, 2.38; 95% CI, 0.90-6.28; P = 0.079). However, presence of *IDH2 p.R140Q* at diagnosis tended to be of favorable prognostic value (HR, 0.18; 95% CI, 0.03-1.37; P

= 0.098). In bone marrow samples at CR, *NRAS p.G12D* was associated significantly with poor OS (HR, 11.58; 95% CI, 1.29-104.04; $P = 0.029$).

Mutational characteristics of PCDHA, NRAS, IDH2, TRRAP, and GATA2

The mutational characteristics of the 5 potential oncogenic drivers or prognostic biomarkers were further investigated (Figure 6). Two *NRAS* mutations, *NRAS p.G12D* and *NRAS p.G13D*, were identified in 12 of 155 diagnosis samples (7.7%) and were mutually exclusive. Fourteen of 155 (9.0%) specimens collected at diagnosis had 1 of 2 variants in the *IDH2* gene: *IDH2 p.R172K* or *IDH2 p.R140K*. These mutations also occurred in a mutually exclusive fashion. Interestingly, both samples harboring *TRRAP p.S722F* corresponded to acute promyelocytic leukemia (APL). The *GATA2 p.R330L* mutation was found only in normal karyotype M2 AML samples.

Validation of TRRAP p.S722F as a co-occurring driver mutation in APL

The presence of *TRRAP p.S722F* was validated with conventional Sanger sequencing (Figure 7). To test the hypothesis that *TRRAP p.S722F* acted as a co-occurring oncogenic driver mutation with a PML/RARα fusion, 24 bone

marrow samples from patients with APL were further analyzed by TaqMan sequencing. Three of these (12.5%) were found to carry *TRRAP p.S722F*. In contrast, results of additional testing for *TRRAP p.S722F* in 32 non-APL samples all were negative. Similarly, the presence of *GATA2 p.R330L* was validated by conventional Sanger sequencing, and 138 bone marrow samples from patients with AML were sequenced via TaqMan. However, no additional *GATA2 p.R330L* mutations were observed, demonstrating that *GATA2 p.R330L* is a rare mutation.

Discussion

Herein, we described recurrent novel mutations in AML, including *PCDHA1 p.L666F*, *SERAC1 p.R645C*, *C7orf50 p.H113Y*, *TRRAP p.S722F*, *FAM178B p.T105M*, *GATA2 p.R330L,* and *PIWIL3 p.N872S*. Results of our in silico prediction using the SIFT/CADD algorithm indicated that some of these mutations are highly likely to be functional. These potentially functional genetic changes included *GATA2 p.R330L* and *TRRAP p.S722F*.

Our primary interest in undertaking this study was to identify clinically relevant, novel SNVs. Our targeted resequencing method focused on 590 small genomic lesions around loci identified via WES. This strategy revealed 7 recurrent loci, and findings from our in silico prediction suggested that 2 of these loci were be highly likely to be functional. Recurrence of a mutation in a specific genetic locus is a strong sign that the mutation confers an oncogenic gain of function. As expected, the frequency of these mutations was consistently below 10%, with a frequency of 4.6% in *PCDHA1 p.L666F*, 6.2% in *SERAC1 p.R645C*, 1.3% in *GATA2 p.R330L*, 2.6% in *C7orf50 p.H113Y*, 1.3% in *TRRAP p.S722F*, and 2.6% in *PIWIL3 p.N872S* in samples collected at diagnosis.

It is somewhat unexpected that the *GATA2* gene is a putative oncogene in

AML. *GATA2* haplo-insufficiency has been known to be related to immune defects,[20] and this gene may be thought to function as a tumor suppressor gene. However, recurrent mutations in the same locus of the *GATA2* gene—as found in our study—support the presumption that the *GATA2* gene itself could function as a putative oncogene in AML. In fact, a loss-of-function mutation in *GATA2* has been reported to be a germline mutation.[21] In the current study, we validated and confirmed that *GATA2 p.R330L* is a somatic mutation, which suggests a disparate role from that of a germline mutation in AML. The proposed role of *GATA2* as a transcription factor[22] further supports the hypothesis that *GATA2* is a potential oncogene in AML. A putative gain-of-function mutation in *GATA2* in hematologic malignancies also has been suggested.[23] An open database shows that amplification of *GATA2* is frequently found in various cancers, including lymphoma, esophagus, prostate, and breast cancer (http://www.cbioportal.org). The very same position locus *GATA2 c.989 or p.330* is the most recurrently altered locus in colon cancer where GATA2 is mutated in more than 1% of samples. From functional perspective, *GATA2 c.989 or p.330* is a locus in small linker area between two GATA zinc fingers of GATA2 protein, suggesting high functionality of mutation in this region. It should be noted that however, the interpretation should be cautious because mutation locus in transcription factor could be varied even if the mutation is oncogenic. In summary, we

assert that *GATA2 p.R330L* has not been previously reported and is a somatic oncogenic mutation, rather than a hereditary germline mutation, as previously noted in AML.[24]

TRRAP p.S722F was the most intriguing mutation found in our study. Although *TRRAP*, which encodes a chromatin modifier,[25] functions as a potential oncogene in various cancers,[26, 27] the role of *TRRAP* in leukemia has not been well characterized. Surprisingly, the same mutation (*TRRAP p.S722F*) has been detected in melanoma,[28] and its putative role as an oncogene was suggested in that study. This mutation was found predominantly in APL, the hallmark of which is the PML-RARα fusion. It is well established that a single oncogenic mutation is insufficient for carcinogenesis, and cooperating additional mutations are necessary for cancer development.[29] The BCR-ABL fusion in acute lymphoblastic leukemia is a good example of this thesis in the hematology field, where *PAX5*, *IKZF1*, *CDKN2A*, and *CDKN2B* cooperate as secondary mutations.[30] In this context, we presume that *TRRAP p.S722F* is an oncogenic mutation that cooperates with the PML-RARα fusion in AML. We also validated, in a separate APL cohort, that 12.5% of APL patients harbor *TRRAP p.S722F*, which supports our assumption. Interestingly, 1 patient with *TRRAP p.S722F* also harbored *NRAS p.G12D*, which implies

that these 2 mutations were cooperative. When the clinical course of this patient with *TRRAP p.S722F* was reviewed, the patient was determined to be in CR after treatment with an all-trans-retinoic acid (ATRA)–based agent, suggesting that this mutation is not related to ATRA resistance.

The mutational pattern of *C7orf50 p.H113Y* supports its role as a potential oncogenic mutation. However, little is known about the functionality of *C7orf50*, and we did not investigate the occurrence of this mutation in a large cohort. Functionality of this mutation should be evaluated in future studies.

We obtained samples at various clinical stages, which enabled us to evaluate the clinical significance of specific mutations. We determined that the *PCDHA1 p.L666F* mutation may be associated with adverse outcomes in AML. This mutation was found frequently in specimens corresponding to CR, persistence, and diagnosis, and the presence of *PCDHA1 p.L666F* was related to unfavorable RFS. We were unable to determine the prognostic significance of *GATA2 p.R330L* or *TRRAP P.S722F*. Given the rarity of these mutations, the prognostic significance should be probed in a large population. The fact that the clinical significance of TCGA–reported mutations in genes such as *IDH2*, *DNMT3A*, and *NRAS* were in concordance with a recent landscape report supports the adequacy of clinical

interpretation in our study.[3]

The sequencing scheme used in our study prevented an analysis of tumor suppressor genes. In addition, the genomic landscape of AML in the Korean population could not be examined comprehensively in the current study because our targeted sequencing was limited to genetic regions found in 53 WES-processed samples. Nevertheless, we maintain that genomic deciphering of cancer is closely related to therapeutics, and drugs targeting relatively frequent oncogenic mutations are being developed. Similar to *FLT3*[8], *IDH1*, and *IDH2*[9] genes, we expect that further functional evaluation of *GATA2*, *TRRAP*, and *C7orf50* will inform strategies to treat patients with cancer who harbor these mutations.

In conclusion, we identified 7 novel mutations in AML. *GATA2 p.R330L* and *TRRAP p.S722F* are oncogenic mutations that, to our knowledge, have not been described previously in AML. *TRRAP p.S722F* is particularly intriguing because it appears to occur as a partner mutation of the PML-RARα fusion in more than 10% of patients with APL.

References

1. Döhner, H., D.J. Weisdorf, and C.D. Bloomfield, *Acute Myeloid Leukemia.* New England Journal of Medicine, 2015. **373**(12): p. 1136-1152.
2. Dombret, H. and C. Gardin, *An update of current treatments for adult acute myeloid leukemia.* Blood, 2016. **127**(1): p. 53-61.
3. Papaemmanuil, E., et al., *Genomic Classification and Prognosis in Acute Myeloid Leukemia.* N Engl J Med, 2016. **374**(23): p. 2209-21.
4. Ilyas, A.M., et al., *Next generation sequencing of acute myeloid leukemia: influencing prognosis.* BMC Genomics, 2015. **16 Suppl 1**: p. S5.
5. Anderson, K., et al., *Genetic variegation of clonal architecture and propagating cells in leukaemia.* Nature, 2011. **469**(7330): p. 356-61.
6. Ding, L., et al., *Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing.* Nature, 2012. **481**(7382): p. 506-10.
7. Wong, T.N., et al., *Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia.* Nature, 2015. **518**(7540): p. 552-5.
8. Cortes, J.E., et al., *Phase I study of quizartinib administered daily to patients with relapsed or refractory acute myeloid leukemia irrespective of FMS-like tyrosine kinase 3-internal tandem duplication status.* J Clin Oncol, 2013. **31**(29): p. 3681-7.
9. Medeiros, B.C., et al., *Isocitrate dehydrogenase mutations in myeloid malignancies.* Leukemia, 2016.
10. Dohner, H., et al., *Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet.* Blood, 2010. **115**(3): p. 453-74.
11. Grimwade, D., et al., *Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult*

patients treated in the United Kingdom Medical Research Council trials. Blood, 2010. **116**(3): p. 354-65.

12. Cheson, B.D., et al., *Revised recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia.* J Clin Oncol, 2003. **21**(24): p. 4642-9.
13. de Greef, G.E., et al., *Criteria for defining a complete remission in acute myeloid leukaemia revisited. An analysis of patients treated in HOVON-SAKK co-operative group studies.* Br J Haematol, 2005. **128**(2): p. 184-91.
14. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.
15. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
16. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.
17. Lee, S., et al., *Korean Variant Archive (KOVA): a reference database of genetic variations in the Korean population.* Sci Rep, 2017. **7**(1): p. 4287.
18. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nat Genet, 2014. **46**(3): p. 310-315.
19. Tavtigian, S.V., et al., *In silico analysis of missense substitutions using sequence-alignment based methods.* Human mutation, 2008. **29**(11): p. 1327-1336.
20. Collin, M., et al., *Human dendritic cell deficiency: the missing ID?* Nat Rev Immunol, 2011. **11**(9): p. 575-83.
21. Crispino, J.D. and M.S. Horwitz, *GATA factor mutations in hematologic disease.* Blood, 2017. **129**(15): p. 2103-2110.
22. Rodriguez-Bravo, V., et al., *The role of GATA2 in lethal prostate cancer aggressiveness.* Nat Rev Urol, 2017. **14**(1): p. 38-48.
23. Mazumder, T.H., A. Uddin, and S. Chakraborty, *Transcription factor gene GATA2: Association of leukemia and nonsynonymous to the synonymous substitution rate across five mammals.* Genomics, 2016. **107**(4): p. 155-

61.
24. Koeffler, H.P. and G. Leong, *Preleukemia: one name, many meanings.* Leukemia, 2017.
25. Ichim, G., et al., *The histone acetyltransferase component TRRAP is targeted for destruction during the cell cycle.* Oncogene, 2014. **33**(2): p. 181-92.
26. Wurdak, H., et al., *An RNAi screen identifies TRRAP as a regulator of brain tumor-initiating cell differentiation.* Cell Stem Cell, 2010. **6**(1): p. 37-47.
27. Mouradov, D., et al., *Colorectal cancer cell lines are representative models of the main molecular subtypes of primary cancer.* Cancer Res, 2014. **74**(12): p. 3238-47.
28. Wei, X., et al., *Exome sequencing identifies GRIN2A as frequently mutated in melanoma.* Nat Genet, 2011. **43**(5): p. 442-6.
29. Vogelstein, B. and K.W. Kinzler, *The Path to Cancer --Three Strikes and You're Out.* N Engl J Med, 2015. **373**(20): p. 1895-8.
30. Moorman, A.V., *New and emerging prognostic and predictive genetic biomarkers in B-cell precursor acute lymphoblastic leukemia.* Haematologica, 2016. **101**(4): p. 407-16.

Figure 1. Searching algorithm for germline and somatic mutations. The 1000 Genomes Project database, the Exsome Aggregation Consortium (ExAC) database, and the OncoPanel were used in the filtering process.

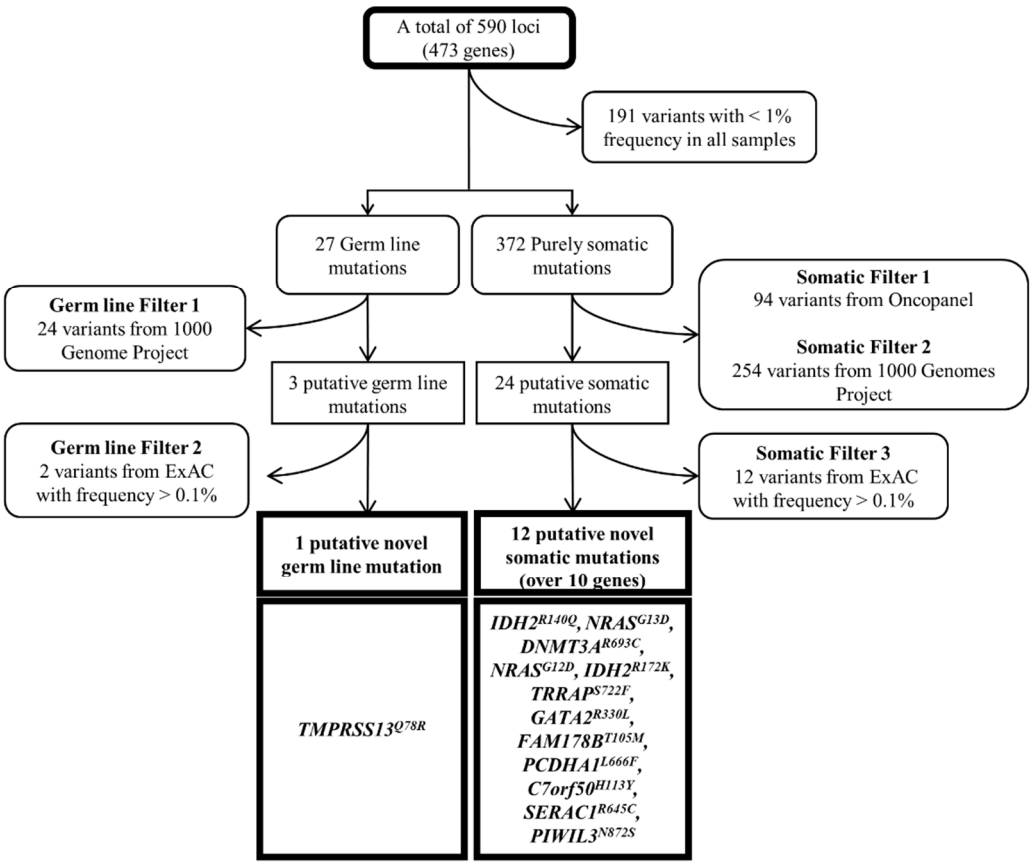


Figure 2. Comparison of mutational frequency in 12 loci between our study and the TCGA database.

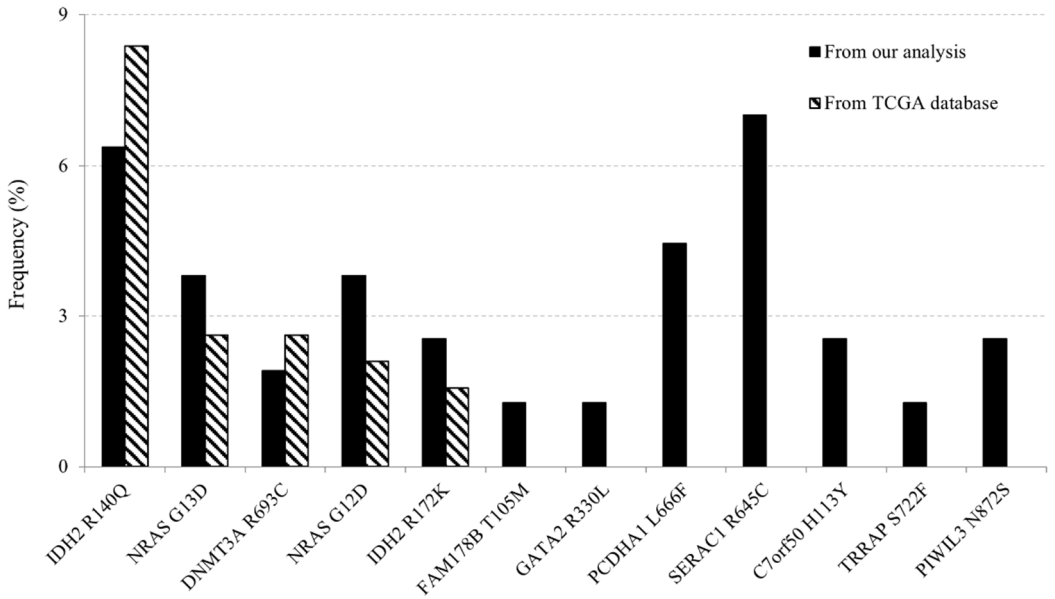


Figure 3. Number of mutations per patient in samples collected at several clinical statuses.

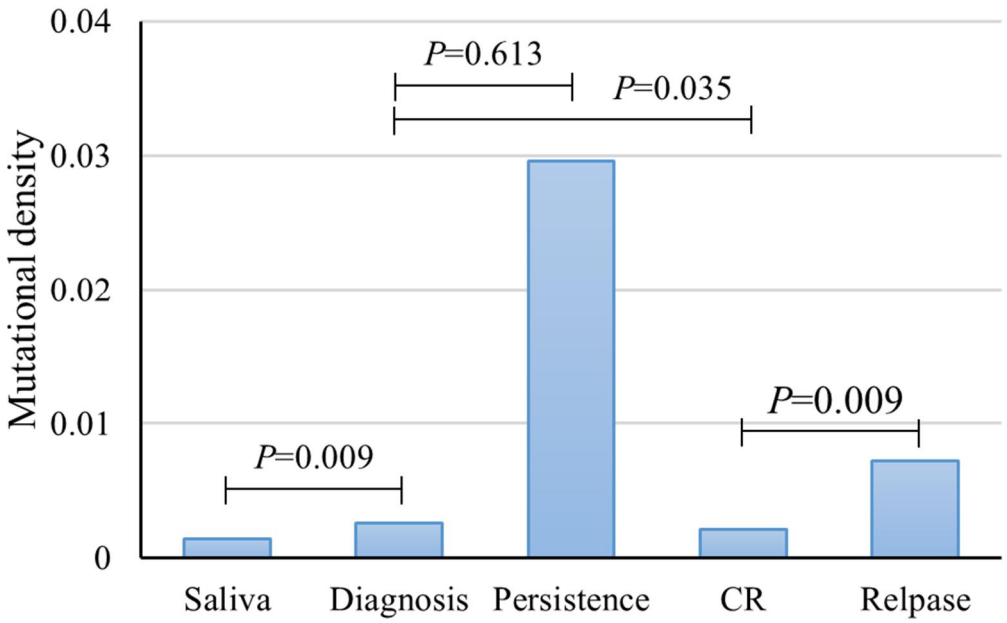


Figure 4. Frequency of novel putative oncogenic mutations in AML samples obtained at various clinical stages (A) and their overlap patterns (B, C) The circos plot demonstrates correlations between pairs of mutations. The width of the arches indicates the percentage of the mutations. (B) The correlations between pairs of mutations are shown. Each tile in this figure denotes $-\log_{10}(p\text{-values})$, where p-values were estimated by Fisher's exact test (C).

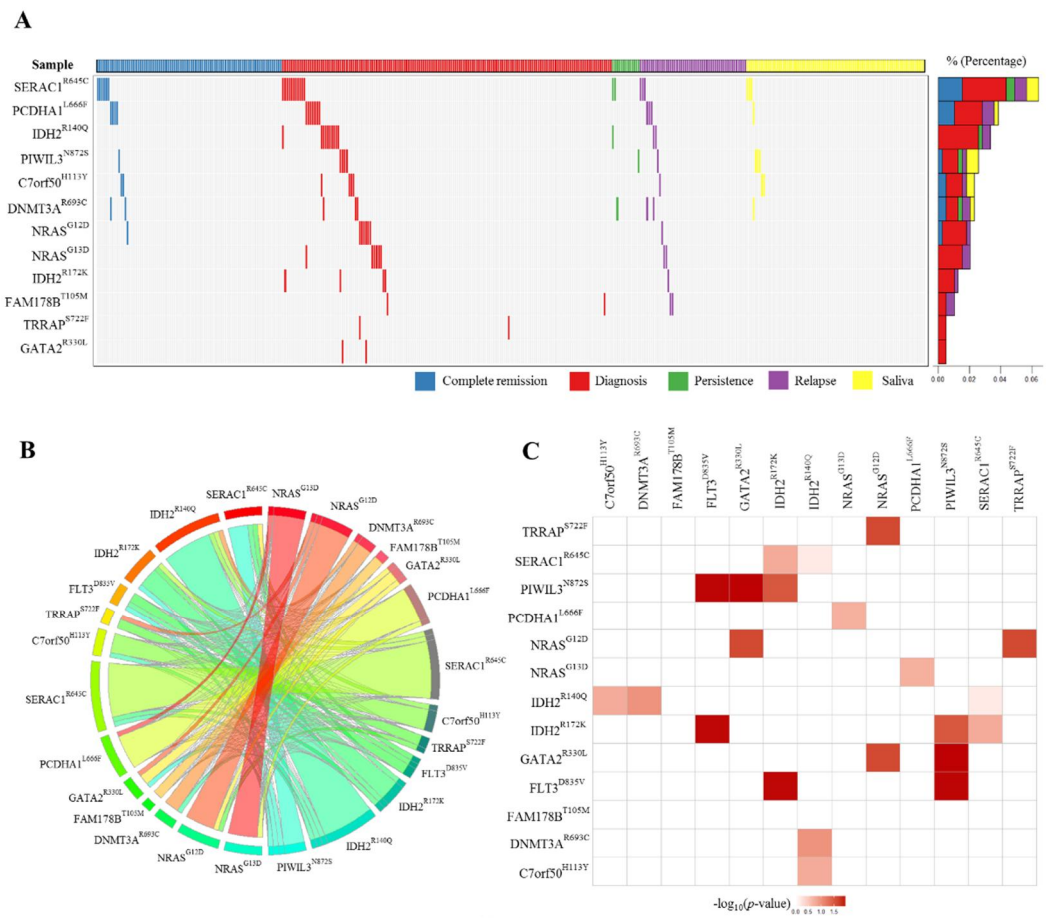
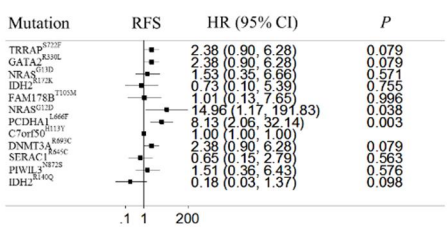
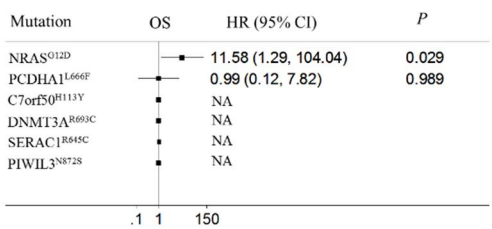


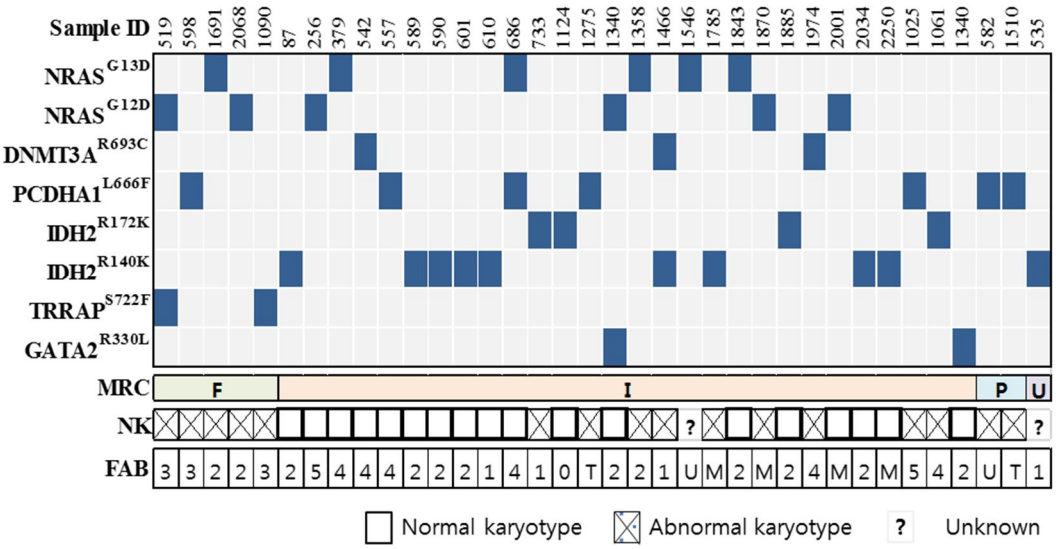
Figure 5. Forest plots summarizing hazard ratios (HRs) of 12 identified mutations, with 95% confidence intervals (CIs) and P-values. (A) Relapse-free survival (RFS) of mutations in diagnosis samples, and (B) overall survival (OS) of mutations in complete remission samples. A multivariate Cox proportional hazard regression model was applied. HRs for RFS were adjusted for secondary acute myelogenous leukemia (AML). HRs for OS were adjusted for age, sex, MRC classification, and secondary AML.





Abbreviations: NA, not available.

Figure 6. Mutational status of *NRAS*, *DNMT3A*, *PCDHA1*, *IDH2*, *TRRAP* and *GATA2* in samples collected from patients at diagnosis. Blue boxes indicate mutations.



Abbreviations: FAB, French-America-British; MRC, Medical Research Council; NK, normal karyotype; F, favorable risk group; I, intermediate risk group; P, poor risk group.

Figure 7. Sanger validation of the *TRRAP p.S722F* mutation. Forward and reverse primer sequences were 5'-GAATGTGTGATACAGGCTTGG-3' and 5'-AAAAAGCCAGGCACACTCAC-3', respectively.

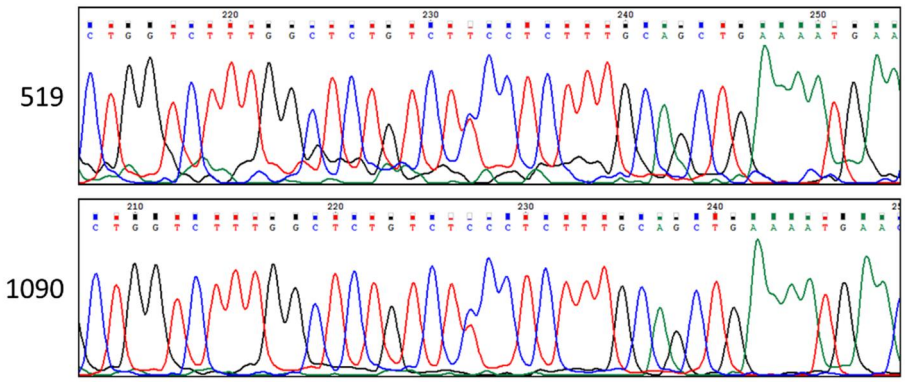


Figure 8. Variant allele frequency of putative oncogenic variants

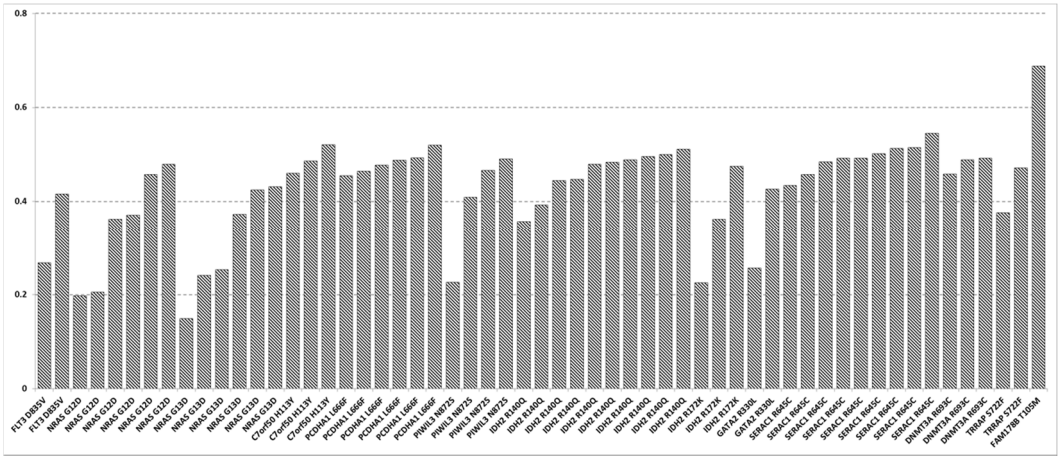


Table 1. Types of pair allele fraction (PAF) used in calling method Adiscan

Allele direction	Normal allele depths		Tumor allele depths		PAF type
	A	B	A’	B’’	
Opposite direction	0	0	25	0	1
	25	0	24	25	2
	25	24	0	25	3
	25	0	0	25	4
	25	24	24	25	5
Same direction	24	25	0	25	6
	25	0	25	24	7
	24	25	24	25	8

Table 2. Clinical characteristics of 389 bone marrow and saliva samples.

Values	Total (N=389)	Saliva (N=84)	Diagnosis (N=155)	CR (N=87)	Relapse (N=50)	Persistence (N=13)	P-value
Sex, n(%)							0.870‡
Male	157 (40.4)	50 (60.0)	97 (62.6)	49 (56.3)	29 (58.0)	7 (53.8)	
Female	232 (59.6)	34 (40.0)	58 (37.4)	38 (43.7)	21 (42.0)	6 (46.2)	
Age, mean (years)	52	48	52	49	55	53	0.022†
Secondary AML, n(%)	58 (14.9)	5 (6.0)	26 (16.7)	14 (16.1)	10 (20.0)	3 (23.1)	0.071‡
MRC stratification, n(%)							0.001‡
Favorable	97 (24.9)	32 (38.0)	34 (21.9)	23 (26.0)	7 (14.0)	1 (7.7)	
Intermediate	243 (62.5)	44 (52.0)	99 (63.9)	54 (62.0)	37 (74.0)	9 (69.0)	
Poor	33 (8.5)	2 (2.4)	15 (9.7)	8 (9.2)	5 (10.0)	3 (23.0)	
Unknown	16 (4.1)	6 (7.1)	7 (4.5)	2 (2.3)	1 (2.0)	0 (0.0)	
FAB classification, n(%)							0.073‡
0	6 (1.5)	0 (0.0)	2 (1.3)	1 (1.1)	3 (3.4)	0 (0.0)	
1	36 (9.2)	9 (10.7)	13 (8.4)	8 (9.2)	6 (6.9)	0 (0.0)	
2	116 (29.8)	24 (28.6)	45 (29.0)	24 (27.6)	18 (20.7)	5 (38.5)	

3	49 (12.6)	20 (23.8)	16 (10.3)	11 (12.6)	1 (1.1)	1 (7.7)	
4	61 (15.7)	15 (17.9)	22 (14.2)	16 (18.4)	6 (6.9)	2 (15.4)	
5	28 (7.2)	3 (3.6)	16 (10.3)	6 (6.9)	1 (1.1)	2 (15.4)	
6	17 (4.4)	4 (4.8)	9 (5.8)	3 (3.4)	0 (0.0)	1 (7.7)	
7	4 (1.0)	0 (0.0)	2 (1.3)	0 (0.0)	2 (2.3)	0 (0.0)	
BMT, n(%)	117 (30.1)	19 (22.6)	45 (29.0)	36 (41.4)	14 (28.0)	3 (23.1)	0.120‡

† By Kruskal-wallis test

‡ By Fisher's exact test

Abbreviations: CR, complete remission; AML, acute myelogenous leukemia; MRC, Medical Research Council; FAB, French-America-British; n, number.

Table 3. List of genetic loci and changes included in targeted resequencing

Gene	Chr	Accession	Nucleic acid change	Type	KoVA	ExAC
APC	5	NM_000038	c.G4479T	Missense	0.000	0.000
C3	19	NM_000064	c.G2827A	Missense	0.000	0.000
CP	3	NM_000096	c.A1493G	Missense	0.011	0.001
JAK3	19	NM_000215	c.G2164A	Missense	0.000	0.008
JAK3	19	NM_000215	c.C1715T	Missense	0.000	0.000
KIT	4	NM_000222	c.G154A	Missense	0.000	0.000
KIT	4	NM_000222	c.A1621C	Missense	0.000	0.077
KIT	4	NM_000222	c.T1669A	Missense	0.000	0.000
KIT	4	NM_000222	c.G1675A	Missense	0.000	0.000
KIT	4	NM_000222	c.T1676A	Missense	0.000	0.000
KIT	4	NM_000222	c.T1679A	Missense	0.000	0.000
KIT	4	NM_000222	c.T1727C	Missense	0.000	0.000
KIT	4	NM_000222	c.C1755T	synonymous	0.000	0.000
KIT	4	NM_000222	c.A1924G	Missense	0.000	0.000
KIT	4	NM_000222	c.T1961C	Missense	0.000	0.000
KIT	4	NM_000222	c.C2009T	Missense	0.000	0.000
KIT	4	NM_000222	c.G2446T	Missense	0.000	0.000
KIT	4	NM_000222	c.A2447T	Missense	0.000	0.000
KIT	4	NM_000222	c.T2466A	Missense	0.000	0.000
KIT	4	NM_000222	c.T2467G	Missense	0.000	0.000
KIT	4	NM_000222	c.T2474C	Missense	0.000	0.000

KIT	4	NM_000222	c.G2515A	Missense	0.000	0.000
MYOC	1	NM_000261	c.C1008G	Missense	0.000	0.000
NF1	17	NM_000267	c.C4084G	Missense	0.000	0.000
PAX2	10	NM_000278	c.A1058C	Missense	0.000	0.000
TPI1	12	NM_000365	c.A517C	Missense	0.000	0.000
KEL	7	NM_000420	c.T2C	Missense	0.000	0.001
STK11	19	NM_000455	c.G165C	Missense	0.000	0.000
STK11	19	NM_000455	c.C842A	Missense	0.000	0.000
STK11	19	NM_000455	c.C1062G	Missense	0.000	0.006
TSC2	16	NM_000548	c.C3475T	Missense	0.000	0.000
TAP1	6	NM_000593	c.G2123A	Missense	0.006	0.029
ORM2	9	NM_000608	c.G328A	Missense	0.000	0.001
CHRNB4	15	NM_000750	c.G946A	Missense	0.000	0.000
CYP1A2	15	NM_000761	c.C1313T	Missense	0.000	0.002
PGR	11	NM_000926	c.A2149T	Missense	0.000	0.000
ARHGEF37	5	NM_001001669	c.C1756A	Missense	0.352	0.164
OR5D18	11	NM_001001952	c.G368A	Missense	0.000	0.000
OR5A2	11	NM_001001954	c.C588A	Missense	0.000	0.000
NT5C3A	7	NM_001002009	c.G22A	Missense	0.000	0.000
TSPYL6	2	NM_001003937	c.G955A	Missense	0.000	0.000
C1orf168	1	NM_001004303	c.C721G	Missense	0.018	0.002
ZNF774	15	NM_001004309	c.A1358G	Missense	0.000	0.008
PLEKHG7	12	NM_001004330	c.A470T	Missense	0.000	0.001
OR10J5	1	NM_001004469	c.C36A	Missense	0.000	0.005
OR4N5	14	NM_001004724	c.C364T	Missense	0.000	0.000

OR6Q1	11	NM_001005186	c.G568A	Missense	0.000	0.002
OR4C12	11	NM_001005270	c.A397G	Missense	0.000	0.000
CACNB4	2	NM_001005747	c.C1184T	Missense	0.000	0.000
WDR33	2	NM_001006622	c.G821A	Missense	0.000	0.083
PIWIL3	22	NM_001008496	c.A2615G	Missense	0.000	0.000
LOC401052	3	NM_001008737	c.C62T	Missense	0.000	0.003
FNIP1	5	NM_001008738	c.G2695A	Missense	0.000	0.001
HRNR	1	NM_001009931	c.C562A	Missense	0.000	0.000
ARSI	5	NM_001012301	c.C1502A	Missense	0.000	0.000
ALG10B	12	NM_001013620	c.A1283C	Missense	0.000	0.000
ALG10B	12	NM_001013620	c.A1283T	Missense	0.000	0.000
SMCO3	12	NM_001013698	c.T145C	Missense	0.000	0.032
FOXB2	9	NM_001013735	c.G904A	Missense	0.000	0.001
TMEM225	11	NM_001013743	c.G565A	Missense	0.015	0.061
MAD1L1	7	NM_001013837	c.A479G	Missense	0.011	0.000
SUGP2	19	NM_001017392	c.T1063C	Missense	0.000	0.000
FANCD2	3	NM_001018115	c.A577G	Missense	0.000	0.007
KCTD21	11	NM_001029859	c.A649G	Missense	0.000	0.001
VSTM4	10	NM_001031746	c.G949A	Missense	0.000	0.000
PHKB	16	NM_001031835	c.G25A	Missense	0.000	0.001
MTUS2	13	NM_001033602	c.A1250T	Missense	0.000	0.001
RYR2	1	NM_001035	c.G4219A	Missense	0.000	0.000
VWA5B1	1	NM_001039500	c.A439C	Missense	0.013	0.000
MUC17	7	NM_001040105	c.C8551T	Missense	0.000	0.002
ENGASE	17	NM_001042573	c.C1523T	Missense	0.000	0.002

TMEM8B	9	NM_001042590	c.C1162T	Missense	0.000	0.001
SHPRH	6	NM_001042683	c.C2632T	Missense	0.000	0.000
DFNB59	2	NM_001042702	c.T182C	Missense	0.000	0.000
ZNF557	19	NM_001044387	c.G548A	Missense	0.015	0.001
MUTYH	1	NM_001048173	c.850-2A>G	Splicing	0.011	0.001
SSTR3	22	NM_001051	c.C1034T	Missense	0.012	0.000
XIRP2	2	NM_001079810	c.A2788G	Missense	0.000	0.000
CNGA3	2	NM_001079878	c.C1252T	Missense	0.000	0.000
KIAA0368	9	NM_001080398	c.C5991G	Missense	0.000	0.000
MYO5B	18	NM_001080467	c.C2881T	Missense	0.000	0.000
TCTN1	12	NM_001082538	c.G1396T	Missense	0.000	0.023
CASC1	12	NM_001082972	c.G233C	Missense	0.000	0.001
ABCF1	6	NM_001090	c.C398T	Missense	0.000	0.000
ADGRF5	6	NM_001098518	c.C1811T	Missense	0.993	0.904
ZBTB34	9	NM_001099270	c.G973T	Missense	0.020	0.001
DZANK1	20	NM_001099407	c.G1081A	Missense	0.019	0.013
RIMS2	8	NM_001100117	c.G353A	Missense	0.000	0.000
C14orf159	14	NM_001102366	c.G479A	Missense	0.000	0.000
TBC1D8	2	NM_001102426	c.C2694A	Missense	0.000	0.000
HECTD4	12	NM_001109662	c.C8480T	Missense	0.000	0.000
MDGA2	14	NM_001113498	c.G104A	Missense	0.000	0.000
CLIC5	6	NM_001114086	c.A299T	Missense	0.020	0.001
NGEF	2	NM_001114090	c.G1370A	Missense	0.000	0.000
C1QC	1	NM_001114101	c.G559A	Missense	0.000	0.002
FAM178B	2	NM_001122646	c.C314T	Missense	0.000	0.000

TENM2	5	NM_001122679	c.G1049A	Missense	0.000	0.000
PRR32	X	NM_001122716	c.G91A	Missense	0.000	0.000
SLC12A3	16	NM_001126108	c.C539A	Missense	0.015	0.000
TP53	17	NM_001126113	c.G31T	Missense	0.000	0.000
MRO	18	NM_001127176	c.G532A	Missense	0.013	0.000
GFI1	1	NM_001127215	c.A1144C	Missense	0.000	0.000
MYRF	11	NM_001127392	c.A2671G	Missense	0.000	0.000
MET	7	NM_001127500	c.A1124C	Missense	0.000	0.000
TEX36	10	NM_001128202	c.G425A	Missense	0.000	0.001
CACNA1D	3	NM_001128839	c.C2853T	Missense	0.000	0.001
CALHM3	10	NM_001129742	c.C173T	Missense	0.000	0.000
CLK3	15	NM_001130028	c.G1593A	Missense	0.000	0.003
DRC3	17	NM_001130090	c.C713T	Missense	0.000	0.000
LEF1	4	NM_001130713	c.C317T	Missense	0.000	0.001
KIAA0922	4	NM_001131007	c.A2003C	Missense	0.000	0.000
SLC4A4	4	NM_001134742	c.G1271A	Missense	0.000	0.003
DAZAP2	12	NM_001136266	c.C361T	Missense	0.000	0.000
RAD21L1	20	NM_001136566	c.A1268C	Missense	0.013	0.064
TTLL11	9	NM_001139442	c.A1787G	Missense	0.008	0.000
FAM205A	9	NM_001141917	c.T416G	Missense	0.000	0.024
SGK1	6	NM_001143676	c.A94T	Missense	0.000	0.001
NFRKB	11	NM_001143835	c.C2705G	Missense	0.000	0.001
EXPH5	11	NM_001144765	c.G3295A	Missense	0.000	0.001
GATA2	3	NM_001145661	c.G989T	Missense	0.000	0.000
GATA2	3	NM_001145661	c.G952A	Missense	0.000	0.000

SLC6A15	12	NM_001146335	c.T1488G	Missense	0.000	0.000
RBM19	12	NM_001146698	c.G1891A	Missense	0.000	0.000
ANK2	4	NM_001148	c.G6650C	Missense	0.000	0.000
PPP6R3	11	NM_001164160	c.G2173C	Missense	0.012	0.001
FLNB	3	NM_001164317	c.G3583A	Missense	0.000	0.001
ZFHX3	16	NM_001164766	c.C1885T	Missense	0.000	0.000
MLH1	3	NM_001167618	c.T428A	Missense	0.000	0.003
SBNO1	12	NM_001167856	c.C119T	Missense	0.000	0.001
ALS2CR11	2	NM_001168221	c.G2062C	Missense	0.000	0.000
ME2	18	NM_001168335	c.A1432G	Missense	0.000	0.002
CUL7	6	NM_001168370	c.T3668C	Missense	0.000	0.001
CUL7	6	NM_001168370	c.C47G	Missense	0.000	0.000
KIAA0319	6	NM_001168376	c.A289C	Missense	0.977	0.916
MAP7D3	X	NM_001173517	c.G418C	Missense	0.000	0.000
SUCLG2	3	NM_001177599	c.G1276A	Missense	0.000	0.000
ARHGEF28	5	NM_001177693	c.A1408G	Missense	0.000	0.003
ATR	3	NM_001184	c.C325T	Missense	0.000	0.001
TNPO3	7	NM_001191028	c.T518G	Missense	0.000	0.000
CDKN2A	9	NM_001195132	c.G205T	Missense	0.000	0.000
CDKN2A	9	NM_001195132	c.G181T	Missense	0.000	0.000
CDKN2A	9	NM_001195132	c.C172T	Missense	0.000	0.000
LRRC72	7	NM_001195280	c.A742T	Missense	0.000	0.000
ZGPAT	20	NM_001195653	c.C183G	Missense	0.988	0.948
ERAP1	5	NM_001198541	c.A895C	Missense	0.006	0.001
NOL4	18	NM_001198549	c.C994T	Missense	0.000	0.000

AKAP2	9	NM_001198656	c.G2501A	Missense	0.000	0.000
DTNA	18	NM_001198938	c.G2062A	Missense	0.000	0.235
KIF16B	20	NM_001199866	c.G2132A	Missense	0.000	0.002
ZNF559	19	NM_001202406	c.C1720T	Missense	0.000	0.001
SELPLG	12	NM_001206609	c.C460G	Missense	0.000	0.000
SELPLG	12	NM_001206609	c.T458C	Missense	0.000	0.000
TMPRSS13	11	NM_001206790	c.A233G	Missense	0.000	0.000
TMPRSS13	11	NM_001206790	c.C230G	Missense	0.000	0.000
DNAH8	6	NM_001206927	c.C5127T	Missense	0.000	0.014
ZNF233	19	NM_001207005	c.T1984A	Missense	0.000	0.202
NXT2	X	NM_001242617	c.C161G	Missense	0.000	0.000
HLA-A	6	NM_001242758	c.A874G	Missense	0.000	0.035
PPP6R2	22	NM_001242900	c.G2419C	Missense	0.014	0.001
PARVB	22	NM_001243386	c.A97G	Missense	0.000	0.004
CLEC3A	16	NM_001244755	c.G151A	Missense	0.000	0.000
COLEC11	2	NM_001255986	c.A578G	Missense	0.000	0.100
CD48	1	NM_001256030	c.G304C	Missense	0.019	0.095
E2F8	11	NM_001256372	c.G97A	Missense	0.000	0.000
EPHX2	8	NM_001256483	c.G263A	Missense	0.018	0.004
MGLL	3	NM_001256585	c.G568A	Missense	0.015	0.003
CLRN1	3	NM_001256819	c.C9A	Missense	0.000	0.001
TTN	2	NM_001256850	c.C90231G	Missense	0.000	0.000
TTN	2	NM_001256850	c.G22790C	Missense	0.000	0.000
ZNF155	19	NM_001260488	c.C113T	Missense	0.000	0.000
ALDH1L1	3	NM_001270365	c.G889C	Missense	0.012	0.001

NEB	2	NM_001271208	c.C5565A	Missense	0.000	0.000
WDR38	9	NM_001276374	c.G241A	Missense	0.020	0.017
WDR38	9	NM_001276374	c.T595G	Missense	0.020	0.017
DNAH11	7	NM_001277115	c.C9402G	Missense	0.000	0.000
DNAH11	7	NM_001277115	c.C10449T	Missense	0.000	0.000
LILRB1	19	NM_001278399	c.G395T	Missense	0.000	0.000
ERMARD	6	NM_001278532	c.T152C	Missense	0.000	0.000
GAS2L1	22	NM_001278730	c.G1220A	Missense	0.000	0.000
AP1G2	14	NM_001282475	c.G946A	Missense	0.000	0.000
SUSD1	9	NM_001282643	c.T572C	Missense	0.000	0.002
SLC25A17	22	NM_001282727	c.C688T	Missense	0.000	0.000
PTCD2	5	NM_001284403	c.C31T	Missense	0.000	0.000
CLCA1	1	NM_001285	c.C193T	Missense	0.984	0.939
GAS8	16	NM_001286205	c.G158C	Missense	0.000	0.002
REV3L	6	NM_001286431	c.G8051A	Missense	0.000	0.001
FAM120A	9	NM_001286722	c.A1613G	Missense	0.000	0.000
IDH2	15	NM_001289910	c.G359A	Missense	0.000	0.000
IDH2	15	NM_001289910	c.G263A	Missense	0.000	0.000
WDR87	19	NM_001291088	c.A6562G	Missense	0.000	0.000
SCN10A	3	NM_001293306	c.T1157G	Missense	0.000	0.001
ZC3H18	16	NM_001294340	c.G982T	Missense	0.000	0.000
ADGRL2	1	NM_001297705	c.G2939T	Missense	0.000	0.000
SV2C	5	NM_001297716	c.C1095T	Missense	0.000	0.004
RASGEF1B	4	NM_001300735	c.T778A	Missense	0.000	0.000
KMT5B	11	NM_001300908	c.C875T	Missense	0.000	0.001

PPP4C	16	NM_001303506	c.C469G	Missense	0.000	0.000
PTEN	10	NM_001304717	c.T1242C	Missense	0.000	0.000
LMO7	13	NM_001306080	c.G649C	Missense	0.000	0.012
JAKMIP1	4	NM_001306133	c.C1135T	Missense	0.000	0.000
CLDN7	17	NM_001307	c.A622G	Missense	0.000	0.000
ARID3B	15	NM_001307939	c.G502C	Missense	0.012	0.002
MYLK3	16	NM_001308301	c.G280C	Missense	0.000	0.001
ADGRG3	16	NM_001308360	c.G928A	Missense	0.007	0.001
ANO3	11	NM_001313726	c.C2864T	Missense	0.000	0.000
THSD7B	2	NM_001316349	c.C4333G	Missense	0.000	0.000
THSD7B	2	NM_001316349	c.C4333T	Missense	0.000	0.000
LZIC	1	NM_001316973	c.G311A	Missense	0.000	0.001
DAB2	5	NM_001343	c.C886T	Missense	0.000	0.003
CELSR3	3	NM_001407	c.C8335T	Missense	0.000	0.000
GDI1	X	NM_001493	c.G533A	Missense	0.000	0.000
GDI1	X	NM_001493	c.G533C	Missense	0.000	0.000
HCRT	17	NM_001524	c.G200T	Missense	0.000	0.000
IL10RA	11	NM_001558	c.G697A	Missense	0.000	0.002
CFB	6	NM_001710	c.T26A	Missense	0.000	0.039
BHMT	5	NM_001713	c.G555A	Missense	0.000	0.000
BHMT	5	NM_001713	c.G555T	Missense	0.000	0.000
BRDT	1	NM_001726	c.C1230G	Missense	0.000	0.209
CD38	4	NM_001775	c.C418T	Missense	0.016	0.001
COL11A1	1	NM_001854	c.C3455T	Missense	0.000	0.000
CSNK2A1	20	NM_001895	c.G346A	Missense	0.000	0.000

FLG	1	NM_002016	c.C4867T	Missense	0.004	0.001
FOXN2	2	NM_002158	c.G1222A	Missense	0.000	0.000
IGHMBP2	11	NM_002180	c.G1825A	Missense	0.000	0.000
ITPR3	6	NM_002224	c.C6802T	Missense	0.000	0.000
ITPR3	6	NM_002224	c.C6924G	Missense	0.000	0.000
MRC1	10	NM_002438	c.G896A	Missense	0.000	0.000
MSH5	6	NM_002441	c.C826T	Missense	0.000	0.000
MYH8	17	NM_002472	c.C3874T	Missense	0.006	0.001
NFX1	9	NM_002504	c.C1584G	Missense	0.000	0.000
NRAS	1	NM_002524	c.A183T	Missense	0.000	0.000
NRAS	1	NM_002524	c.A182C	Missense	0.000	0.000
NRAS	1	NM_002524	c.C181A	Missense	0.000	0.000
NRAS	1	NM_002524	c.G52A	Missense	0.000	0.000
NRAS	1	NM_002524	c.G38A	Missense	0.000	0.000
NRAS	1	NM_002524	c.G37C	Missense	0.000	0.000
NRAS	1	NM_002524	c.G35A	Missense	0.000	0.000
NRAS	1	NM_002524	c.G34C	Missense	0.000	0.000
PPP2R3A	3	NM_002718	c.A1924G	Missense	0.000	0.263
PSAP	10	NM_002778	c.T337C	Missense	0.000	0.000
PTPRG	3	NM_002841	c.C1541T	Missense	0.000	0.000
GRK1	13	NM_002929	c.A299G	Missense	0.000	0.000
RPS6KA1	1	NM_002953	c.A2101C	Missense	0.000	0.000
S100A3	1	NM_002960	c.G8A	Missense	0.012	0.084
SLC18A2	10	NM_003054	c.A322G	Missense	0.000	0.000
SMARCB1	22	NM_003073	c.G1143A	Missense	0.000	0.000

TFAM	10	NM_003201	c.T326G	Missense	0.000	0.040
TJP1	15	NM_003257	c.G3230A	Missense	0.000	0.000
MZF1	19	NM_003422	c.G308A	Missense	0.014	0.002
ZNF132	19	NM_003433	c.G967T	Missense	0.000	0.001
KMT2D	12	NM_003482	c.T908C	Missense	0.000	0.000
KMT2D	12	NM_003482	c.C907A	Missense	0.000	0.000
KMT2D	12	NM_003482	c.T902G	Missense	0.000	0.000
TRRAP	7	NM_003496	c.C2165T	Missense	0.000	0.000
FZD9	7	NM_003508	c.G1027A	Missense	0.000	0.000
HIST1H3B	6	NM_003537	c.A83T	Missense	0.000	0.000
EEA1	12	NM_003566	c.G3991A	Missense	0.000	0.000
CILP	15	NM_003613	c.C1291T	Missense	0.000	0.001
MCM3AP	21	NM_003906	c.T5609G	Missense	0.026	0.003
FLT3	13	NM_004119	c.A2525G	Missense	0.000	0.000
FLT3	13	NM_004119	c.2520_2520 delinsCGGA TCC,	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c.C2508G	Missense	0.000	0.000
FLT3	13	NM_004119	c.T2505A	Missense	0.000	0.000
FLT3	13	NM_004119	c.A2504T	Missense	0.000	0.000
FLT3	13	NM_004119	c.G2503T	Missense	0.000	0.000
FLT3	13	NM_004119	c.G2501A	Missense	0.000	0.000
FLT3	13	NM_004119	c.G2492A	Missense	0.000	0.000
FLT3	13	NM_004119	c.C2039T	Missense	0.000	0.000
FLT3	13	NM_004119	c.1811_1811 delinsACCG AGAATATG AATATGAT	in/del (inframe)	0.000	0.000

			CTCAAATGGGA,			
FLT3	13	NM_004119	c.1807_1807delinsTATGAATATGATCTCAAAT,	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c.1803_1803delinsCCAGCTACAGATGGTACAGGTGACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTCAGAGAATATGAATATGATCTC,	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c.1798_1798delinsGCAGATAATGAGTACTTCTACGTTGATTTCAGAGAATATGAATATG,	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c.A1796T	Missense	0.000	0.000
FLT3	13	NM_004119	c.T1775C	Missense	0.000	0.000
FLT3	13	NM_004119	c.A1715G	Missense	0.000	0.000
FLT3	13	NM_004119	c.C1352T	Missense	0.000	0.000
AKAP6	14	NM_004274	c.A416G	Missense	0.000	0.001
NFE2L3	7	NM_004289	c.A1850C	Missense	0.000	0.001
CEACAM5	19	NM_004363	c.C2024T	Missense	0.000	0.000
CREBBP	16	NM_004380	c.G6551A	Missense	0.000	0.000

DVL1	1	NM_004421	c.G1955T	Missense	0.000	0.000
ERBB2	17	NM_004448	c.G428A	Missense	0.000	0.001
NOTCH4	6	NM_004557	c.C1753T	Missense	0.000	0.000
PCK2	14	NM_004563	c.C730T	Missense	0.000	0.002
COIL	17	NM_004645	c.G433A	Missense	0.000	0.127
NPHS1	19	NM_004646	c.G1802C	Missense	0.000	0.002
IL27RA	19	NM_004843	c.G1070A	Missense	0.000	0.001
JAK2	9	NM_004972	c.G1849T	Missense	0.000	0.001
MED12	X	NM_005120	c.A2540G	Missense	0.000	0.000
ACTC1	15	NM_005159	c.C634T	Missense	0.000	0.000
ARID3A	19	NM_005224	c.C708A	Missense	0.000	0.000
EPHA3	3	NM_005233	c.G1672C	Missense	0.000	0.000
FAT1	4	NM_005245	c.C7472A	Missense	0.000	0.000
GAK	4	NM_005255	c.A667G	Missense	0.000	0.000
MCHR1	22	NM_005297	c.G1201A	Missense	0.000	0.001
SPRR3	1	NM_005416	c.G289A	Missense	0.000	0.000
APBA2	15	NM_005503	c.C809A	Missense	0.000	0.000
HSPG2	1	NM_005529	c.C7438T	Missense	0.000	0.001
INSL3	19	NM_005543	c.C148T	Missense	0.000	0.000
LAMA5	20	NM_005560	c.G5387A	Missense	0.000	0.000
IDH1	2	NM_005896	c.G395A	Missense	0.000	0.000
IDH1	2	NM_005896	c.C394T	Missense	0.000	0.000
MLLT6	17	NM_005937	c.C1007T	Missense	0.000	0.000
MYH1	17	NM_005963	c.C5077T	Missense	0.000	0.002
SEMA3A	7	NM_006080	c.T307G	Missense	0.000	0.000

YAP1	11	NM_006106	c.C1069T	Missense	0.000	0.000
PCSK5	9	NM_006200	c.C1497G	Missense	0.000	0.000
PDGFRA	4	NM_006206	c.C2472T	synonymous	0.114	0.186
RAD21	8	NM_006265	c.T1238G	Missense	0.000	0.000
TNFAIP2	14	NM_006291	c.C1919T	Missense	0.000	0.001
VARS	6	NM_006295	c.A1868G	Missense	0.012	0.015
SMC1A	X	NM_006306	c.G2369A	Missense	0.000	0.000
PARP4	13	NM_006437	c.C3322T	Missense	0.988	0.931
ZNF234	19	NM_006630	c.G368A	Missense	0.000	0.001
WNK2	9	NM_006648	c.G5350A	Missense	0.000	0.001
SRCAP	16	NM_006662	c.C8851G	Missense	0.000	0.001
UPK2	11	NM_006760	c.C149A	Missense	0.000	0.001
UPK2	11	NM_006760	c.C454T	Missense	0.008	0.000
KRT38	17	NM_006771	c.G139A	Missense	0.000	0.000
GPR75	2	NM_006794	c.C1372T	Missense	0.000	0.000
RIPK3	14	NM_006871	c.T1042C	Missense	0.000	0.001
RXRG	1	NM_006917	c.T230G	Missense	0.000	0.000
TOPBP1	3	NM_007027	c.G4372T	Missense	0.000	0.002
ADAMTS5	21	NM_007038	c.G254A	Missense	0.000	0.000
SLC22A18AS	11	NM_007105	c.C112T	Missense	0.000	0.002
TMF1	3	NM_007114	c.G2267A	Missense	0.000	0.001
TRPA1	8	NM_007332	c.A558C	Missense	0.985	0.710
ELL2	5	NM_012081	c.T984A	Missense	0.014	0.001
FOXB1	15	NM_012182	c.G494T	Missense	0.015	0.002

KIAA1549L	11	NM_012194	c.G3926A	Missense	0.000	0.000
HAVCR1	5	NM_012206	c.T536C	Missense	0.000	0.889
XRN2	20	NM_012255	c.A2494G	Missense	0.000	0.001
SF3B1	2	NM_012433	c.G1998C	Missense	0.000	0.000
PSD4	2	NM_012455	c.C1780T	Missense	0.000	0.001
FLRT3	20	NM_013281	c.A1380C	Missense	0.000	0.031
ADAMTS7	15	NM_014272	c.C3733G	Missense	0.000	0.000
TRAM1	8	NM_014294	c.T857G	Missense	0.000	0.001
HEBP2	6	NM_014320	c.G419A	Missense	0.000	0.001
AFF4	5	NM_014423	c.A554T	Missense	0.000	0.005
BRINP1	9	NM_014618	c.G326A	Missense	0.000	0.000
KIAA0430	16	NM_014647	c.G2593C	Missense	0.000	0.009
ZNF536	19	NM_014717	c.C3113G	Missense	0.000	0.000
FAM65B	6	NM_014722	c.G1112A	Missense	0.000	0.018
KMT2B	19	NM_014727	c.C1760G	Missense	0.012	0.001
ATP2C2	16	NM_014861	c.G2755A	Missense	0.000	0.000
DCLRE1A	10	NM_014881	c.G212A	Missense	0.000	0.003
DIS3	13	NM_014953	c.C1571T	Missense	0.000	0.001
PDCD11	10	NM_014976	c.G2131A	Missense	0.000	0.000
PDCD11	10	NM_014976	c.G2881A	Missense	0.000	0.002
MYH15	3	NM_014981	c.A2600C	Missense	0.000	0.000
SPEN	1	NM_015001	c.C6719T	Missense	0.000	0.000
HECW1	7	NM_015052	c.C2579T	Missense	0.001	0.000
CUL9	6	NM_015089	c.A7130G	Missense	0.000	0.000
ZZEF1	17	NM_015113	c.C7568G	Missense	0.000	0.002

CLCC1	1	NM_015127	c.C1328T	Missense	0.000	0.002
EXOC6B	2	NM_015189	c.C2016A	Missense	0.000	0.000
PLCB1	20	NM_015192	c.T2854C	Missense	0.000	0.000
SMCHD1	18	NM_015295	c.C5126T	Missense	0.000	0.000
KHNYN	14	NM_015299	c.G1220T	Missense	0.012	0.001
HAUS5	19	NM_015302	c.C178T	Missense	0.000	0.002
ICE1	5	NM_015325	c.C1431G	Missense	0.000	0.001
VPS13D	1	NM_015378	c.A8833G	Missense	0.000	0.000
PTPN23	3	NM_015466	c.C3956T	Missense	0.000	0.005
DNAH1	3	NM_015512	c.G5838A	Missense	0.000	0.000
C2CD3	11	NM_015531	c.C131A	Missense	0.000	0.002
DST	6	NM_015548	c.T98C	Missense	0.000	0.060
LRIT1	10	NM_015613	c.A1771G	Missense	0.015	0.046
RGS22	8	NM_015668	c.G1190T	Missense	0.000	0.000
CPAMD8	19	NM_015692	c.G1460A	Missense	0.000	0.000
RDH8	19	NM_015725	c.G666A	Missense	0.000	0.102
FAHD2A	2	NM_016044	c.G154A	Missense	0.000	0.001
ACTL6B	7	NM_016188	c.C609A	Missense	0.000	0.000
CYB5R4	6	NM_016230	c.A1336G	Missense	0.000	0.001
SRRM2	16	NM_016333	c.C7070A	Missense	0.000	0.000
CENPF	1	NM_016343	c.A5642T	Missense	0.000	0.000
DRICH1	22	NM_016449	c.A603T	Missense	0.000	0.003
SPTBN5	15	NM_016642	c.G8700C	Missense	0.000	0.000
NUP54	4	NM_017426	c.G733A	Missense	0.000	0.000
MYH2	17	NM_017534	c.G4583A	Missense	0.000	0.000

DNAH3	16	NM_017539	c.T3308C	Missense	0.000	0.001
NOTCH1	9	NM_017617	c.T5033C	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.T4799C	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.G4793C	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.T4778C	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.T4754C	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.G4732A	Missense	0.000	0.000
NOTCH1	9	NM_017617	c.T4721C	Missense	0.000	0.000
BNC2	9	NM_017637	c.A692G	Missense	0.000	0.000
TMEM104	17	NM_017728	c.G1315A	Missense	0.014	0.002
FOCAD	9	NM_017794	c.A376G	Missense	0.000	0.001
DUS2	16	NM_017803	c.G1405C	Missense	0.000	0.001
TMCO3	13	NM_017905	c.C1451T	Missense	0.000	0.000
ATG2B	14	NM_018036	c.A3266G	Missense	0.000	0.001
ANO1	11	NM_018043	c.T1496G	Missense	0.000	0.000
ASXL2	2	NM_018263	c.C1471T	Missense	0.000	0.000
MBD5	2	NM_018328	c.A3310G	Missense	0.000	0.001
CMTR2	16	NM_018348	c.A1567G	Missense	0.000	0.000
HJURP	2	NM_018410	c.G2192A	Missense	0.000	0.000
UBAP2	9	NM_018449	c.G1874A	Missense	0.000	0.007
LRP1B	2	NM_018557	c.C3867A	Missense	0.000	0.000
PCDHA1	5	NM_018900	c.C1996T	Missense	0.000	0.000
PCDHGA10	5	NM_018913	c.C2072T	Missense	0.000	0.001
PCDHGB6	5	NM_018926	c.A1969G	Missense	0.003	0.000
ANKIB1	7	NM_019004	c.A1952G	Missense	0.000	0.001

DDX56	7	NM_019082	c.A746G	Missense	0.000	0.010
EPB41L4B	9	NM_019114	c.C2131T	Missense	0.000	0.000
KIAA1217	10	NM_019590	c.G5617A	Missense	0.000	0.000
GRIPAP1	X	NM_020137	c.T1934G	Missense	0.000	0.000
YAE1D1	7	NM_020192	c.C379T	Missense	0.000	0.000
YAE1D1	7	NM_020192	c.G388T	Missense	0.000	0.000
MRPL1	4	NM_020236	c.C848G	Missense	0.000	0.001
TRIM49	11	NM_020358	c.C1181A	Missense	0.000	0.000
TRIM49	11	NM_020358	c.G1147A	Missense	0.000	0.000
TRIM49	11	NM_020358	c.C877G	Missense	0.000	0.078
ATP8B2	1	NM_020452	c.G2273A	Missense	0.011	0.001
ATP8B2	1	NM_020452	c.C3347T	Missense	0.000	0.000
UBR4	1	NM_020765	c.A3319G	Missense	0.000	0.143
KIAA1462	10	NM_020848	c.G1205A	Missense	0.000	0.001
KIAA1462	10	NM_020848	c.G334A	Missense	0.000	0.002
DDX55	12	NM_020936	c.G802A	Missense	0.013	0.001
FANCM	14	NM_020937	c.G4931A	Missense	0.020	0.002
WDFY4	10	NM_020945	c.C1746G	Missense	0.000	0.000
AQP9	15	NM_020980	c.T758C	Missense	0.000	0.002
GRIK2	6	NM_021956	c.G1473T	Missense	0.000	0.000
TNN	1	NM_022093	c.T1904G	Missense	0.000	0.000
EDAR	2	NM_022336	c.A1138C	Missense	0.000	0.002
RHBDF1	16	NM_022450	c.C830T	Missense	0.000	0.001
MRPL44	2	NM_022915	c.A578C	Missense	0.000	0.000
SLC27A3	1	NM_024330	c.G1348A	Missense	0.000	0.002

ATP10A	15	NM_024490	c.A1099G	Missense	0.009	0.001
ABHD8	19	NM_024527	c.G121T	Missense	0.000	0.000
ECHDC3	10	NM_024693	c.G856A	Missense	0.000	0.003
TXNDC15	5	NM_024715	c.C594G	Missense	0.000	0.000
PEAK1	15	NM_024776	c.G4459A	Missense	0.000	0.001
PLBD1	12	NM_024829	c.G793A	Missense	0.990	0.955
SPEF2	5	NM_024867	c.G3696C	Missense	0.000	0.000
KIAA0319L	1	NM_024874	c.A700G	Missense	0.000	0.001
PLEKHH3	17	NM_024927	c.G1360C	Missense	0.000	0.001
NYNRIN	14	NM_025081	c.G1399A	Missense	0.000	0.000
CWH43	4	NM_025087	c.T1703C	Missense	0.000	0.000
FBXO38	5	NM_030793	c.C2495A	Missense	0.000	0.000
FAM117A	17	NM_030802	c.T820C	Missense	0.000	0.000
SNX27	1	NM_030918	c.A772G	Missense	0.000	0.000
EPPK1	8	NM_031308	c.G6519C	Missense	0.014	0.001
HNRNPD	4	NM_031369	c.A271G	Missense	0.000	0.000
PCDHA10	5	NM_031859	c.G2012A	Missense	0.000	0.001
HMCN1	1	NM_031935	c.G13576A	Missense	0.000	0.000
KRTAP17-1	17	NM_031964	c.C129A	Missense	0.000	0.000
KRTAP17-1	17	NM_031964	c.G125A	Missense	0.000	0.190
PCDHGB1	5	NM_032095	c.G172C	Missense	0.000	0.003
PSD2	5	NM_032289	c.G2120A	Missense	0.000	0.000
C7orf50	7	NM_032350	c.C337T	Missense	0.000	0.001
SLX4	16	NM_032444	c.T3367C	Missense	0.000	0.000
FNDC1	6	NM_032532	c.G755A	Missense	0.000	0.000

TTYH2	17	NM_032646	c.G541A	Missense	0.000	0.221
BEST3	12	NM_032735	c.T1870A	Missense	0.000	0.000
PHF5A	22	NM_032758	c.G166A	Missense	0.000	0.000
HAVCR2	5	NM_032782	c.A245G	Missense	0.000	0.004
CEP89	19	NM_032816	c.A425C	Missense	0.000	0.395
SERAC1	6	NM_032861	c.C1933T	Missense	0.000	0.001
MICALCL	11	NM_032867	c.G197A	Missense	0.000	0.007
COL27A1	9	NM_032888	c.G152A	Missense	0.000	0.005
COL27A1	9	NM_032888	c.C1163T	Missense	0.000	0.001
BBS4	15	NM_033028	c.A1414G	Missense	0.010	0.001
ADAM19	5	NM_033274	c.T2218G	Missense	0.000	0.156
ADAM19	5	NM_033274	c.G1979A	Missense	0.000	0.098
KRAS	12	NM_033360	c.G436A	Missense	0.000	0.000
KRAS	12	NM_033360	c.A183C	Missense	0.000	0.000
KRAS	12	NM_033360	c.A182C	Missense	0.000	0.000
KRAS	12	NM_033360	c.C181A	Missense	0.000	0.000
KRAS	12	NM_033360	c.G175A	Missense	0.000	0.000
KRAS	12	NM_033360	c.C64A	Missense	0.000	0.000
KRAS	12	NM_033360	c.G57C	Missense	0.000	0.000
KRAS	12	NM_033360	c.G38A	Missense	0.000	0.000
KRAS	12	NM_033360	c.G37T	Missense	0.000	0.000
KRAS	12	NM_033360	c.G35T	Missense	0.000	0.000
KRAS	12	NM_033360	c.G34T	Missense	0.000	0.000
MID2	X	NM_052817	c.C557T	Missense	0.000	0.000
SLC18B1	6	NM_052831	c.A643G	Missense	0.000	0.001

IL17F	6	NM_052872	c.A377G	Missense	0.015	0.067
GPRIN1	5	NM_052899	c.G1157A	Missense	0.000	0.000
SIGLEC12	19	NM_053003	c.A1757G	Missense	0.012	0.015
COL6A3	2	NM_057165	c.G599A	Missense	0.000	0.000
VPS16	20	NM_080413	c.A1394T	Missense	0.000	0.001
RBBP8NL	20	NM_080833	c.A545G	Missense	0.000	0.000
LRPPRC	2	NM_133259	c.A1432G	Missense	0.011	0.006
TTN	2	NM_133379	c.G5645A	Missense	0.000	0.000
VWA5B2	3	NM_138345	c.G1255A	Missense	0.000	0.000
SDSL	12	NM_138432	c.G322C	Missense	0.024	0.002
TMC1	9	NM_138691	c.A439G	Missense	0.000	0.000
TEX30	13	NM_138779	c.A248G	Missense	0.000	0.000
SON	21	NM_138927	c.T1993G	Missense	0.000	0.000
ADAMTS17	15	NM_139057	c.G2023A	Missense	0.000	0.000
CASC5	15	NM_144508	c.G571A	Missense	0.000	0.000
ANKRD22	10	NM_144590	c.A443C	Missense	0.016	0.046
ANKRD35	1	NM_144698	c.T580C	Missense	0.000	0.001
IL23R	1	NM_144701	c.T929C	Missense	0.983	0.879
10-Sep	2	NM_144710	c.G1070A	Missense	0.000	0.000
ZNF485	10	NM_145312	c.C64T	Missense	0.000	0.102
AKAP9	7	NM_147185	c.G5252T	Missense	0.000	0.000
AKAP9	7	NM_147185	c.A8165G	Missense	0.000	0.000
OSBPL9	1	NM_148906	c.C1453G	Missense	0.000	0.000
NXPE1	11	NM_152315	c.G940A	Missense	0.000	0.000
AGBL1	15	NM_152336	c.G98A	Missense	0.000	0.001

MYOM3	1	NM_152372	c.A431T	Missense	0.026	0.007
AHSA2	2	NM_152392	c.G309T	Missense	0.000	0.001
WBSCR27	7	NM_152559	c.C638T	Missense	0.000	0.002
CCDC185	1	NM_152610	c.C628A	Missense	0.000	0.000
GLIS3	9	NM_152629	c.A2006G	Missense	0.000	0.001
TXLNB	6	NM_153235	c.G1042A	Missense	0.000	0.153
C6orf223	6	NM_153246	c.A671C	Missense	0.000	0.000
LRIG3	12	NM_153377	c.C1843T	Missense	0.000	0.000
LRIG3	12	NM_153377	c.G302A	Missense	0.000	0.001
COG7	16	NM_153603	c.T2103C	Missense	0.031	0.058
COG7	16	NM_153603	c.C1814T	Missense	0.026	0.053
CMYA5	5	NM_153610	c.G7144T	Missense	0.000	0.002
KCNJ1	11	NM_153767	c.A199G	Missense	0.000	0.019
KCNH1	1	NM_172362	c.C856T	Missense	0.000	0.000
TRIM65	17	NM_173547	c.C448T	Missense	0.012	0.000
TRIM65	17	NM_173547	c.A446T	Missense	0.012	0.000
FAM217A	6	NM_173563	c.T85A	Missense	0.000	0.000
DNAH17	17	NM_173628	c.C7330T	Missense	0.000	0.000
FSIP2	2	NM_173651	c.G18715A	Missense	0.000	0.000
C8orf31	8	NM_173687	c.A395G	Missense	0.000	0.000
LVRN	5	NM_173800	c.G189C	Missense	0.024	0.004
IL1RN	2	NM_173842	c.G370A	Missense	0.000	0.003
N6AMT2	13	NM_174928	c.G232T	Missense	0.000	0.000
DNMT3A	2	NM_175629	c.C2644T	Missense	0.000	0.000
DNMT3A	2	NM_175629	c.T1901A	Missense	0.000	0.000

C12orf60	12	NM_175874	c.G702T	Missense	0.000	0.001
PPFIA1	11	NM_177423	c.G2284A	Missense	0.000	0.001
TREML1	6	NM_178174	c.C710G	Missense	0.000	0.000
LHFPL1	X	NM_178175	c.C365T	Missense	0.000	0.000
PRR30	2	NM_178553	c.C271T	Missense	0.021	0.003
TRIML1	4	NM_178556	c.C387A	Missense	0.016	0.001
NLRC3	16	NM_178844	c.G2306A	Missense	0.000	0.026
PRR5-ARHGAP8	22	NM_181334	c.G914A	Missense	0.022	0.001
KRTAP19-5	21	NM_181611	c.G98A	Missense	0.000	0.000
BPIFB4	20	NM_182519	c.C241T	Missense	0.000	0.001
YIPF7	4	NM_182592	c.C337A	Missense	0.000	0.000
ACSM2B	16	NM_182617	c.C404A	Missense	0.000	0.000
GEN1	2	NM_182625	c.C2692T	Missense	0.000	0.007
BPIFB3	20	NM_182658	c.G178T	Missense	0.012	0.001
FAM19A3	1	NM_182759	c.G41A	Missense	0.000	0.000
MACC1	7	NM_182762	c.A1438T	Missense	0.000	0.001
P4HA3	11	NM_182904	c.C1334T	Missense	0.004	0.001
TMPRSS9	19	NM_182973	c.A3148G	Missense	0.013	0.000
DST	6	NM_183380	c.G13262A	Missense	0.000	0.001
B3GLCT	13	NM_194318	c.G1330A	Missense	0.000	0.000
CCDC57	17	NM_198082	c.A519T	Missense	0.000	0.003
CSMD3	8	NM_198124	c.G2414A	Missense	0.000	0.000
PRSS55	8	NM_198464	c.G755A	Missense	0.000	0.000
NRK	X	NM_198465	c.C1222T	Missense	0.000	0.000

ZNF615	19	NM_198480	c.G1259A	Missense	0.000	0.001
SULF2	20	NM_198596	c.G19A	Missense	0.000	0.000
TDRD1	10	NM_198795	c.G281A	Missense	0.000	0.001
SERPINB8	18	NM_198833	c.G823A	Missense	0.000	0.001
CNNM2	10	NM_199076	c.G604A	Missense	0.000	0.001
FAM179A	2	NM_199280	c.C1709A	Missense	0.000	0.000
AFTPH	2	NM_203437	c.G2519A	Missense	0.000	0.002
USH2A	1	NM_206933	c.C7294T	Missense	0.000	0.000
TRPM3	9	NM_206945	c.C4235A	Missense	0.000	0.000
ZNF474	5	NM_207317	c.C757T	Missense	0.008	0.001
GADL1	3	NM_207359	c.G617T	Missense	0.000	0.001
NCKAP5	2	NM_207363	c.C1667T	Missense	0.000	0.000
ADAMTSL3	15	NM_207517	c.C1141T	Missense	0.000	0.001
NRAS	1	NM_002524	c182_183AA_TG	Missense	0.000	0.000
NRAS	1	NM_002524	c181_183CAA_AAG	Missense	0.000	0.000
NRAS	1	NM_002524	c181_182CA_AG	Missense	0.000	0.000
NRAS	1	NM_002524	c180_181AC_TA	Missense	0.000	0.000
NRAS	1	NM_002524	c38_39GT_TC	Missense	0.000	0.000
KRAS	12	NM_002524	c180_181TC_CA	Missense	0.000	0.000
KRAS	12	NM_002524	c38_39GC_TG	Missense	0.000	0.000
KRAS	12	NM_002524	c37_39GGC_CGT	Missense	0.000	0.000

KRAS	12	NM_002524	c36_37TG_AT	Missense	0.000	0.000
KRAS	12	NM_002524	c35_36GT_TC	Missense	0.000	0.000
KRAS	12	NM_002524	c34_36GGT_TGC	Missense	0.000	0.000
FLT3	13	NM_004119	c2509_2510AT_CC	Missense	0.000	0.000
FLT3	13	NM_004119	c2508_2510delCAT	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c2506_2508delATC	in/del (inframe)	0.000	0.000
FLT3	13	NM_004119	c2503_2505delGAT	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1669_1674delTGGAAG	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1669_1683del15	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1672_1680del9	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1673_1687del15	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1675_1677delGTT	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c1735_1737delGAT	in/del (inframe)	0.000	0.000
KIT	4	NM_000222	c2143_2145delAGC	in/del (inframe)	0.000	0.000
CDKN2A	9	NM_001195132	c171_172CC_TT	Missense	0.000	0.000
JAK2	9	NM_004972	c1848_1849TG_CT	Missense	0.000	0.000

Table 4. Twelve somatic mutations discovered in our filtering process The functional consequences of each variant were predicted using the SIFT/CADD algorithm. A SIFT score ≤0.05 was taken as indicative of deleterious substitutions.

Gene	Chr	Locus	Ref	Alt	Amino acid change	Accession	Type	SIFT	CADD
NRAS	1	115258744	C	T	G13D	NM_002524	missense	0.03	5.048
NRAS	1	115258747	C	T	G12D	NM_002524	missense	0	5.524
DNMT3A	2	25457243	G	A	R693C	NM_153759	missense	0	4.224
FAM178B	2	97637888	G	A	T105M	NM_001122646	missense	0.19	-1.094
GATA2	3	128202731	C	A	R330L	NM_001145661	missense	0	5.528
PCDHA1	5	140167871	C	T	L666F	NM_018900	missense	0.01	0.581
SERAC1	6	158532430	G	A	R645C	NM_032861	missense	0.26	2.217
C7orf50	7	1040174	G	A	H113Y	NM_001134395	missense	0.04	3.132
TRRAP	7	98509802	C	T	S722F	NM_003496	missense	0	4.926
IDH2	15	90631838	C	T	R172K	NM_002168	missense	0	2.957
IDH2	15	90631934	C	T	R140Q	NM_002168	missense	0	3.604
PIWIL3	22	25115473	T	C	N872S	NM_001008496	missense	1	-0.668

Abbreviations: Chr., chromosome; Ref., reference sequence; Alt., altered sequence; SIFT, Sorting Intolerant From Tolerant; CADD, Combined Annotation Dependent Depletion

Table 5. Univariate Cox proportional hazard regression analyses of clinical factors for overall survival (OS) and relapse-free survival (RFS) in patients with samples (A) at diagnosis, and (B) in complete remission.

(A) For diagnosis samples

		RFS		OS	
Factors	**References**	**HR (95% CI)**	***P*-value**	**HR (95% CI)**	***P*-value**
Clinical factors					
Age	**> 52 (vs. ≤ 52)**	1.24 (0.68-2.30)	0.474	1.92 (1.09-3.41)	**0.025**
Sex	**Male (vs. female)**	1.54 (0.81-2.95)	0.191	1.98 (1.08-3.66)	**0.028**
MRC	**Unfavorable**	2.93 (0.85-10.14)	0.090	4.07 (1.92-8.62)	**< 0.001**
Secondary AML	**Yes (vs. no)**	**2.28 (0.99-5.26)**	**0.054**	2.33 (1.26-4.32)	**0.007**
Allo-HSCT	**Yes (vs. no)**	1.13 (0.61-2.08)	0.696	0.67 (0.37-1.23)	0.198
Somatic mutations					
***TRRAP*^{S722F}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***FLT3*^{D835V}**	**Mutant (vs. wild type)**	0.61 (0.08-4.50)	0.630	NA	NA
***GATA2*^{R330L}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***NRAS*^{G13D}**	**Mutant (vs. wild type)**	1.53 (0.37-6.42)	0.560	1.01 (0.24-4.13)	0.994
***IDH2*^{R172K}**	**Mutant (vs. wild type)**	0.61 (0.08-4.50)	0.630	NA	NA
***FAM178B*^{T105M}**	**Mutant (vs. wild type)**	0.91 (0.12-6.86)	0.928	0.98 (0.13-7.17)	0.980
***NRAS*^{G12D}**	**Mutant (vs. wild type)**	**21.49 (1.95-238.07)**	**0.012**	0.42 (0.06-3.03)	0.389
***PCDHA1*^{L666F}**	**Mutant (vs. wild type)**	**6.38 (1.69-24.13)**	**0.006**	2.00 (0.62-6.47)	0.245
***C7orf50*^{H113Y}**	**Mutant (vs. wild type)**	NA	NA	NA	
***DNMT3A*^{R693C}**	**Mutant (vs. wild type)**	NA	NA	2.80 (0.68-11.54)	0.153
***SERAC1*^{R645C}**	**Mutant (vs. wild type)**	0.72 (0.17-2.99)	0.649	1.06 (0.38-2.94)	0.910
***PIWIL3*^{N872S}**	**Mutant (vs. wild type)**	1.25 (0.30-5.24)	0.760	0.59 (0.08-4.28)	0.603
***IDH2*^{R140Q}**	**Mutant (vs. wild type)**	0.43 (0.13-1.38)	0.156	1.17 (0.42-3.25)	0.761
Germline mutation					
***TMPRSS13*^{Q78R}**	**Mutant (vs. wild type)**	0.67 (0.08-5.34)	0.705	NA	NA

(B) For CR samples

Factors	References	RFS		OS	
		HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Age	**> 52 (vs. ≤ 52)**	1.47 (0.71-3.03)	0.295	8.35 (1.89-36.90)	**0.005**
Sex	**Male (vs. Female)**	0.84 (0.40-1.73)	0.632	2.63 (0.85-8.15)	0.095
MRC	**Unfavorable**	4.67 (0.56-38.84)	0.153	5.50 (1.74-17.39)	**0.003**
Secondary AML	**Yes (vs. No)**	1.73 (0.59-5.06)	0.320	1.58 (0.45-5.59)	0.475
Allo-HSCT	**Yes (vs. No)**	1.02 (0.50-2.10)	0.954	0.81 (0.29-2.22)	0.675
Somatic mutations					
***TRRAP*^{S722F}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***FLT3*^{D835V}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***GATA2*^{R330L}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***NRAS*^{G13D}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***IDH2*^{R172K}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***FAM178B*^{T105M}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***NRAS*^{G12D}**	**Mutant (vs. wild type)**	NA	NA	**20.55 (2.30-183.83)**	**0.007**
***PCDHA1*^{L666F}**	**Mutant (vs. wild type)**	0.40 (0.05-2.97)	0.368	1.76 (0.23-13.35)	0.585
***C7orf50*^{H113Y}**	**Mutant (vs. wild type)**	0.28 (0.04-2.08)	0.214	NA	NA
***DNMT3A*^{R693C}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
***SERAC1*^{R645C}**	**Mutant (vs. wild type)**	1.02 (0.30-3.51)	0.975	NA	NA
***PIWIL3*^{N872S}**	**Mutant (vs. wild type)**	0.53 (0.07-3.98)	0.540	NA	NA
***IDH2*^{R140Q}**	**Mutant (vs. wild type)**	NA	NA	NA	NA
Germline mutation					
***TMPRSS13*^{Q78R}**	**Mutant (vs. wild type)**	7.48 (0.47-119.82)	0.155	NA	NA

Abbreviations: MRC, Medical Research Council; AML, acute myelogenous leukemia; Allo-HSCT, allogeneic hematopoietic stem cell transplantation; HR, hazard ratio.

국문 초록

연구 배경

차세대 유전체 시퀀싱 기법의 발전에 힘입어 급성 골수성 백혈병의 유전변이에 관한 많은 부분들이 밝혀졌다. 그러나, 아직도 그 빈도가 높지 않은 유전변이들이 다 밝혀지지는 아니하였으며, 추가 연구를 통하여 밝힐 가치가 있다. 본 연구에서는 차세대 유전체 시퀀싱 기법으로 다양한 임상적 상태의 골수시료를 이용하여 급성 골수성 백혈병의 돌연변이를 분석하고 그 임상적 의미를 밝혀보려 하였다.

연구 방법

본 연구에서는 53 명의 한국인 급성 골수성 백혈병 환자의 샘플로 전장 엑솜시퀀싱 (whole exome sequencing, WES) 을 수행하고, 이어 389 명 급성 골수성 백혈병 환자의 샘플로 타겟 리시퀀싱을 수행하였다. 타겟 리시퀀싱은 WES 에서 발견된 모든 아미노산 변화를 수반하는 변이에 대해 수행되었다. 타겟 리시퀀싱에 사용된 샘플은 진단시 샘플 150개, 재발/불응성 샘플 63개 그리고 완전관해시 샘플 87개였다. 데이터의 분석을 위하여 Exome Aggregation Consortium (ExAC), 1000 Genome Database, Korean Variant Archive (KOVA), 그리고 The Cancer Genome Atlas (TCGA) 의 데이터베이스를 참고하였다.

연구 결과

전장 엑솜 시퀀싱의 결과 590 개의 위치에서 돌연변이가 발견되어, 이 590개

의 위치에 대한 타겟 리시퀀싱이 이루어졌다. 타겟 리시퀀싱 결과, 29개의 변이들이 1% 이상의 샘플들에서 반복적으로 발견되었다. 이들 중 이미 잘 알려진 암돌연변이 *IDH2 p.R172K, IDH2 p.R140H, NRAS p.G13D, NRAS p.G12D, DNMT3A p.R693C* 의 빈도는 TCGA 와 크게 다르지 않았다. 한편, 총 일곱 개의 유전변이 (*PCDHA1 p.L666F*, *SERAC1 p.R645C*, *C7orf50 p.H113Y*, *TRRAP p.S722F*, *FAM178 p.T105M*, *GATA2 p.R330L*, *PIWIL3 p.N872S*) 는 이전 문헌보고에서 알려지지 않은 급성 골수성 백혈병의 새로운 중요 유전변이로 생각되었다. 특히 *TRRAP p.S722F* 변이와 *GATA2 p.R330L* 변이는 인실리코 예상 도구 이용 시 기능성이 높게 측정되었다. 흑색종에서 암돌연변이로 기보고된 *TRRAP p.S722F* 는 추가 검증에서 급성 전골수성 백혈병 환자의 12.8% 에서 발견되었고, 급성 전골수성 이와의 급성 골수성 백혈병 환자에서는 발견되지 않았다.

결론

본 연구를 통하여 한국인 급성 골수성 백혈병 환자들의 돌연변이 분포도가 TCGA 와 크게 다르지 않음을 확인하였다. 그리고 *TRRAP p.S722F* 를 포함한 새로운 7개의 유전변이를 보고하는 바이다. 특히 *TRRAP p.S722F*, 은 급성 전골수성 백혈병의 10% 이상 환자군에서 핵심 유전변이인 PML-RARα 융합돌연변이의 보조인자로 작용하는 것으로 보인다.

주요어: 급성 골수성 백혈병, 급성 전골수성 백혈병, TRRAP, 차세대 유전체 시퀀싱

학번: 2010-30558